

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 13:07:33 ; Search time 233 Seconds
(without alignments)
2240.734 Million cell updates/sec

Title: US-10-659-055-3
Perfect score: 4020
Sequence: 1 ADPGSHHHHSHKRYTTLT.....AHQHIYTHSHFIKQCFSLP 740

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3939	98.0	766	1 DPP4_HUMAN	P27487 h dipeptidyl
2	3939	98.0	766	2 Q53TN1_HUMAN	Q53tn1 homo sapien
3	3910	97.3	765	2 Q5R7G7_PONPY	Q5r7g7 pongo pygma
4	3585	89.2	765	1 DPP4_BOVIN	P81425 b dipeptidyl
5	3560	88.6	765	1 DPP4_FELCA	Q9n217 f dipeptidyl
6	3548	88.3	766	1 DPP4_PIG	P22411 s dipeptidyl
7	3408.5	84.8	767	1 DPP4_RAT	P14740 r dipeptidyl
8	3390	84.3	760	1 DPP4_MOUSE	P28843 m dipeptidyl
9	2520.5	62.7	751	2 Q75S83_AGGHB	Q75s83 agkistrodon
10	2519.5	62.7	751	2 Q75S82_AGGHB	Q75s82 agkistrodon
11	2519	62.7	759	2 Q5Z1B1_CHICK	Q5zi81 gallus gall
12	2348	58.4	737	2 Q641D6_XENLA	Q641d6 xenopus lae
13	2331.5	58.0	748	2 P70092_XENLA	P70092 xenopus lae
14	2175	54.1	760	2 Q53TP5_HUMAN	Q53tp5 homo sapien
15	2168	53.9	760	1 SEPR_HUMAN	Q12884 homo sapien
16	2160.5	53.7	761	2 Q8R452_RAT	Q8r492 rattus norv
17	2158.5	53.7	761	1 SEPR_MOUSE	P97321 mus musculu
18	2115	52.6	755	2 Q91651_XENLA	Q91651 xenopus lae
19	1922.5	49.6	860	2 Q4RQJ4_TETNG	Q4rqj4 tetraodon n
20	1819.5	45.3	704	2 Q4S309_TETNG	Q4s309 tetraodon n
21	1222	30.4	796	1 DPP10_HUMAN	Q8n608 homo sapien
22	1210	30.1	796	1 DPP10_RAT	Q6g629 rattus norv
23	1202.5	29.9	796	2 Q6INB7_XENLA	Q6inb7 xenopus lae
24	1183	29.4	797	1 DPP10_MOUSE	Q6nxx7 mus musculu
25	1129	28.1	865	1 DPP6_HUMAN	P42658 homo sapien
26	1124	28.0	731	2 Q7PSF9_ANOGA	Q7psf9 anopheles g
27	1121	27.9	803	1 DPP6_PANTR	Q5is50 pan troglod
28	1118	27.8	803	2 Q9QV78_9MURI	Q9qv78 rattus sp.
29	1116	27.8	859	1 DPP6_RAT	P46101 rattus norv
30	1112.5	27.7	745	2 Q9VMM2_DROME	Q9vmm2 drosophila
31	1106	27.5	863	1 DPP6_BOVIN	P42659 bos taurus

32	1089	27.1	801	2 Q5U4C2_MOUSE	Q5u4c2 mus musculu
33	1089	27.1	803	2 Q80VM5_MOUSE	Q80vm5 mus musculu
34	1089	27.1	804	1 DPP6_MOUSE	Q92218 mus musculu
35	1088.5	27.1	439	2 Q6PG25_XENLA	Q6pg25 xenopus lae
36	1069.5	26.6	815	2 Q4RK63_TETNG	Q4rk63 tetraodon n
37	1059	26.3	724	2 Q5IS95_9PRIM	Q5is95 saimiri bol
38	1051	26.1	604	2 Q53SS4_HUMAN	Q53ss4 homo sapien
39	1045	26.0	801	2 Q4RK66_TETNG	Q4rk66 tetraodon n
40	1038	25.8	802	2 Q9VUH1_DROME	Q9vuh1 drosophila
41	1025.5	25.5	935	2 Q3VMB4_DROME	Q3vmb4 drosophila
42	998	24.8	775	2 Q5J6J3_TIRIU	Q5j6j3 trichophyco
43	996	24.8	842	2 Q4RVF8_TETNG	Q4rvf8 tetraodon n
44	994	24.7	874	2 Q7SI80_EMENI	Q7si80 emericella
45	986	24.5	880	2 Q5B934_EMENI	Q5b934 aspergillus

ALIGNMENTS

RESULT 1
DPP4_HUMAN STANDARD; PRT; 766 AA.
AC P27487;
DT 01-AUG-1992 (Rel. 23, Created)
DT 13-FEB-1996 (Rel. 33, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DB Dipeptidyl peptidase 4 (EC 3.4.14.5) (dipeptidyl peptidase IV) (DPP
DB IV) (T-cell activation antigen CD26) (TP103) (Adenosine deaminase
DB complexing protein 2) (ADAPF) [Contains: dipeptidyl peptidase 4
DB membrane form (dipeptidyl peptidase IV membrane form); dipeptidyl
DB peptidase 4 soluble form (dipeptidyl peptidase IV soluble form)].
GN Name=DPP4; Synonyms=ADCP2, CD26;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=92329551; PubMed=1352704; DOI=10.1016/0167-4781(92)90036-Y;
RA Misumi Y., Hayashi Y., Arakawa F., Ikehara Y.;
RT "Molecular cloning and sequence analysis of human dipeptidyl peptidase
IV, a serine proteinase on the cell surface.";
RL Blochim. Biophys. Acta 1131:333-336(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RX MEDLINE=95012454; PubMed=7927537;
RA Abbott C.A., Baker E., Sutherland G.R., McCaughan G.W.;
RT "Genomic organization, exact localization, and tissue expression of
the human CD26 (dipeptidyl peptidase IV) gene.";
RL Immunogenetics 40:331-338(1994).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheal blood;
RX MEDLINE=92325476; PubMed=1352530;
RA Tanaka T., Camerini D., Seed B., Torimoto Y., Dang N.H., Kameoka J.,
Dahlberg H.N., Schlossman S.F., Morimoto C.;
RT "Cloning and functional expression of the T cell activation antigen
CD26.";
RL J. Immunol. 149:481-486(1992).
RN [4]
RP ERRATUM.
RX MEDLINE=93171637; PubMed=8094732;
RA Tanaka T.;
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Prostate, and Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altechul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boesak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skaleka U., Smalius D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 1-551.
RC TISSUE=Colon;
RX MEDLINE=92165847; PubMed=1347043;
RA Darmoul D., Lacasa M., Baricault L., Marguet D., Sapin C., Trotot P.,
RA Barbat A.;
RT "Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like
RT colon cancer cell lines HT-29 and Caco-2. Cloning of the complete
RT human coding sequence and changes of dipeptidyl peptidase IV mRNA
RT levels during cell differentiation.";
RL J. Biol. Chem. 267:4824-4833(1992).
RN [7]
RP NUCLEOTIDE SEQUENCE OF 545-766.
RC TISSUE=Colon;
RX MEDLINE=91024044; PubMed=1977364;
RA Darmoul D., Lacasa M., Chantret I., Swallow D., Trugnan G.;
RT "Isolation of a cDNA probe for the human intestinal
RT dipeptidylpeptidase IV and assignment of the gene locus DPP4 to
RT chromosome 2".
RL Ann. Hum. Genet. 54:191-197(1990).
RN [8]
RP NUCLEOTIDE SEQUENCE OF 1-31.
RX MEDLINE=96067599; PubMed=7487939;
RA Boehm S.K., Gum J.R. Jr., Erickson R.H., Hicks J.W., Kim Y.S.;
RT "Human dipeptidyl peptidase IV gene promoter: tissue-specific
RT regulation from a TATA-less GC-rich sequence characteristic of a
RT housekeeping gene promoter.";
RL Biochem. J. 311:835-843(1995).
RN [9]
RP PROTEIN SEQUENCE OF 1-22, AND TISSUE SPECIFICITY.
RX PubMed=1677636;
RA Gorvel J.P., Ferrero A., Chambrud L., Rigal A., Bonicel J.,
RA Maroux S.;
RT "Expression of sucrase-isomaltase and dipeptidylpeptidase IV in human
RT small intestine and colon.";
RL Gastroenterology 101:618-625(1991).
RN [10]
RP PARTIAL PROTEIN SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=93210468; PubMed=8096237; DOI=10.1084/jem.177.4.1135;
RA Morrison M.E., Vijayasathadi S., Engelstein D., Albino A.P.,
RA Houghton A.N.;
RT "A marker for neoplastic progression of human melanocytes is a cell
RT surface ectopeptidase.";
RL J. Exp. Med. 177:1135-1143(1993).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 38-766 IN COMPLEX WITH
RP INHIBITOR, AND HOMODIMERIZATION.
RX PubMed=12832764; DOI=10.1107/S0907444903010059;
RA Oefner C., D'Arcy A., Mac Sweeney A., Pizarro S., Gardiner R.,
RA Dale G.B.;
RT "High-resolution structure of human apo dipeptidyl peptidase IV/CD26
RT and its complex with 1-[(2-[(5-iodopyridin-2-yl)amino]-ethyl)amino]-
RT acetyl]-2-cyano-(S)-pyrrolidine.";
RL Acta Crystallogr. D 59:1206-1212(2003).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-771, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=12846248; DOI=10.1016/S0006-291X(03)00258-4;
RA Higamatsu H., Kyono K., Higashiyama Y., Fukushima C., Shima H.,
RA Sugiyama S., Inaka K., Yamamoto A., Shimizu R.;
RT "The structure and function of human dipeptidyl peptidase IV,
RT possessing a unique eight-bladed beta-propeller fold.";
RL Biochem. Biophys. Res. Commun. 302:849-854(2003).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=12483204; DOI=10.1038/nb882;
RA Rasmussen H.B., Branner S., Wiberg P.C., Wegmann N.;
RT "Crystal structure of human dipeptidyl peptidase IV/CD26 in complex
RT with a substrate analog.";
RL Nat. Struct. Biol. 10:19-25(2003).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=12906826; DOI=10.1016/S0969-2126(03)00160-6;
RA Thoma R., Loeffler B., Stihle M., Huber W., Ruf A., Hennig M.;
RT "Structural basis of proline-specific exopeptidase activity as
RT observed in human dipeptidyl peptidase-IV.";
RL Structure 11:947-959(2003).
CC -|- FUNCTION: Removes N-terminal dipeptides sequentially from the
CC polypeptides having unsubstituted N-termini provided that the
CC penultimate residue is proline. Plays a role in T cell activation.
CC -|- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -|- SUBUNIT: Homodimer or heterodimer with seprase (PAP).
CC -|- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC soluble form.
CC -|- TISSUE SPECIFICITY: Expressed in the poorly differentiated crypt
CC cells of the small intestine as well as in the mature villous
CC cells. Expressed at very low levels in the colon.
CC -|- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC by proteolytic processing.
CC -|- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC EMBL; U13735; AAB60646.1; -; Genomic DNA.
CC EMBL; U13710; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13711; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13712; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13713; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13714; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13715; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13716; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13717; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13718; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13719; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13720; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13721; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13722; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13723; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13724; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13725; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13726; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13727; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13728; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13729; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13730; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13731; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13732; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13733; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13734; AAB60646.1; JOINED; Genomic DNA.

DR EMBL; M74777; AAA51943.1; -; mRNA.
DR EMBL; BC013329; AAH13329.2; -; mRNA.

Query Match 98.0%; Score 3939; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 2e-259;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTDEF 72
Db |||||
Qy 39 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTDEF 98
Db |||||

Qy 73 GHSINDYSISPDGQFILLEYNVYKQWRHSYTSASDIYDLNKRQLITEERIPNNTQWTS 132
Db |||||

Qy 99 GHSINDYSISPDGQFILLEYNVYKQWRHSYTSASDIYDLNKRQLITEERIPNNTQWTS 158
Db |||||

Qy 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTYVEEVSAYGALWSP 192
Db |||||

Qy 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTYVEEVSAYGALWSP 218
Db |||||

Qy 193 NGTFLAYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFPVNTDLS 252
Db |||||

Qy 219 NGTFLAYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFPVNTDLS 278
Db |||||

Qy 253 VTNATSIQITAPASMLIGDHYLSDVWATQERISLOWLRRIQNSYVMDICDYDESSGRWN 312
Db |||||

Qy 279 VTNATSIQITAPASMLIGDHYLSDVWATQERISLOWLRRIQNSYVMDICDYDESSGRWN 338
Db |||||

Qy 313 CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFTFI 372
Db |||||

Qy 339 CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFTFI 398
Db |||||

Qy 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSYTKVTCLSCELNPERCOYS 432
Db |||||

Qy 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSYTKVTCLSCELNPERCOYS 458
Db |||||

Qy 433 LNETKFWQMLPPHFDKSKYKPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 492
Db |||||

Qy 519 LNETKFWQMLPPHFDKSKYKPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 578
Db |||||

Qy 553 DGRSGYQGDKIMHAINRRLGTPEVDQIEAARQFSKMGFVNDKRIAIWGSYGYVTSM 612
Db |||||

Qy 579 DGRSGYQGDKIMHAINRRLGTPEVDQIEAARQFSKMGFVNDKRIAIWGSYGYVTSM 638
Db |||||

Qy 613 VLGSYGVPKCGIATAVPSRWEYDSTYTERYMGFLTPEDNLDHYRNSTVMSRAENKQV 672
Db |||||

Qy 639 VLGSYGVPKCGIATAVPSRWEYDSTYTERYMGFLTPEDNLDHYRNSTVMSRAENKQV 698
Db |||||

Qy 673 EYLLIHGTADNVHFQSAQISKALVDGVDFQAMWYTDDEHGTASSTAHQHIYTHMSHF 732
Db |||||

Qy 699 EYLLIHGTADNVHFQSAQISKALVDGVDFQAMWYTDDEHGTASSTAHQHIYTHMSHF 758
Db |||||

Qy 733 IKQCFSLP 740
Db |||||

Qy 759 IKQCFSLP 766
Db |||||

RESULT 2
Q53TN1 HUMAN
ID Q53TN1 HUMAN PRELIMINARY; PRT; 766 AA.
AC Q53TN1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein DPP4.
GN Name=DPP4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.

NCBI_TaxID=9606;
[1]
RN NUCLEOTIDE SEQUENCE.
RP Ozanich A., Stoneking T., Hawkins M., Sapetti L.;
RA "the sequence of Homo sapiens BAC clone RP11-178A14.";
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RP Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RP Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RP Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC008063; AA931179.1; -; Genomic_DNA.
DR SMR; Q53TN1; 39-766.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9 AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estra.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SRR; 1.
KW Hydrolase; Hypothetical protein; Transmembrane.
SQ SEQUENCE 766 AA; 88279 MW; 5FB4A2C6662D6117 CRC64;

Query Match 98.0%; Score 3939; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 2e-259;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTDEF 72
Db |||||

Qy 39 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTDEF 98
Db |||||

Qy 73 GHSINDYSISPDGQFILLEYNVYKQWRHSYTSASDIYDLNKRQLITEERIPNNTQWTS 132
Db |||||

Qy 99 GHSINDYSISPDGQFILLEYNVYKQWRHSYTSASDIYDLNKRQLITEERIPNNTQWTS 158
Db |||||

Qy 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTYVEEVSAYGALWSP 192
Db |||||

Qy 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTYVEEVSAYGALWSP 218
Db |||||

Qy 193 NGTFLAYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFPVNTDLS 252
Db |||||

Qy 219 NGTFLAYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFPVNTDLS 278
Db |||||

Qy 253 VTNATSIQITAPASMLIGDHYLSDVWATQERISLOWLRRIQNSYVMDICDYDESSGRWN 312
Db |||||

Qy 279 VTNATSIQITAPASMLIGDHYLSDVWATQERISLOWLRRIQNSYVMDICDYDESSGRWN 338
Db |||||

Qy 313 CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFTFI 372
Db |||||

Qy 339 CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFTFI 398
Db |||||

Qy 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSYTKVTCLSCELNPERCOYS 432
Db |||||

Qy 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSYTKVTCLSCELNPERCOYS 458
Db |||||

Qy 433 VSFSEAKYQLRCSGPGFLPYTLHSSVNDKGLRVLEDNSALDKQLQNVQMPKCLDFII 492
Db |||||

Qy 459 VSFSEAKYQLRCSGPGFLPYTLHSSVNDKGLRVLEDNSALDKQLQNVQMPKCLDFII 518
Db |||||

Qy 493 LNETKFWQMLPPHFDKSKYKPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 552
Db |||||

```

Db 519 LNETKFWQMLPPHFDKSKYPPLLDVYAGPCSKADTVPRLNWATYLASTENIIVASF 578
Qy 553 DGRSGYGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 612
Db 579 DGRSGYGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 638
Qy 613 VLGGSGYGVKCGIAVAPVSRWEYDVSVYTERVMGLPTPEDNLDHYRNSVMSRAENFKQV 672
Db 639 VLGGSGYGVKCGIAVAPVSRWEYDVSVYTERVMGLPTPEDNLDHYRNSVMSRAENFKQV 698
Qy 673 EYLLIHGTADDNVHFQQAQISKALVDVGDFQAMWYTDDEHGIIASSTAHOHIYTHMSHF 732
Db 699 EYLLIHGTADDNVHFQQAQISKALVDVGDFQAMWYTDDEHGIIASSTAHOHIYTHMSHF 758
Qy 733 IKQCFSLP 740
Db 759 IKQCFSLP 765

RESULT 3
QSR7G7 PONPY PRELIMINARY; PRT; 765 AA.
ID QSR7G7;
AC QSR7G7;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp469P1419;
GN Name=DKFZp469P1419;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC The German CNA Consortium;
RA Pouetka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Well B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; CR860150; CAH92293.1; -; mRNA.
DR SMR; QSR7G7; 38-765.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; P:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept S9 AS.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV N.1.
DR PROSITE; PS00708; Peptidase S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP SER; 1.
KW Hydrolase; Hypothetical protein; Transmembrane.
SQ SEQUENCE 765 AA; 88038 MW; 6F81ECD98C3D397 CRC64;

Query Match 97.3%; Score 3910; DB 2; Length 765;
Best Local Similarity 99.2%; Pred. No. 1.9e-257;
Matches 722; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 13 SRKTYTLDTLQNTYRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVFLNSTFDF 72
Db 38 SRKTYTLDTLQNTYRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVFLNSTFDF 97
Qy 73 GHSINDYSTSPGQFTLLRYNKKWRHSYTSYDYLKQQLITEERIPNNTQWTVS 132
Db 98 GHSINDYSTSPGQFTLLRYNKKWRHSYTSYDYLKQQLITEERIPNNTQWTVS 157
Qy 133 PVGHKLAYVWNNDIYVKIEPNLPSRITWTGKEDIYNGITDWVYEEVFSAISALWSP 192
Db 158 PVGHKLAYVWNNDIYVKIEPNLPSRITWTGKEDIYNGITDWVYEEVFSAISALWSP 217

```

RESULT 4

```

DPP4_BOVIN
ID DPP4_BOVIN STANDARD; PRT; 765 AA.
AC P81425; Q8WMSG;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE IV) (T-cell activation antigen CD26) (Adenosine deaminase complexing
DE protein) (ADCP-I) (Activation molecule 3) (ACT3) (WC10) [Contains:
DE Dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane
DE form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV
DE soluble form)].
GN Name=DPP4; Synonyms=CD26;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lymphocyte;
RX MEDLINE=22067734; PubMed=12073152; DOI=10.1007/s00251-002-0456-6;
RA Lee S.-U., Park Y.-H., Davis W.C., Hamilton M.J., Naessens J.,
RA Bohach G.A.;
RT "Molecular characterization of bovine CD26 upregulated by a
RT staphylococcal superantigen.";
RL Immunogenetics 54:216-220 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 1-17.
RC TISSUE=Thymus;

```


RX MEDLINE=22021197; PubMed=11981836;
 RX DOI=10.1002/1521.4141(200205)32:5<1472::AID-IMMU1472>3.0.CO;2-Q;
 RA Gliddon D.R., Howard C.J.;
 RT "CD26 is expressed on a restricted subpopulation of dendritic cells in
 RT vivo";
 RL Eur. J. Immunol. 32:1472-1481(2002).
 RN [3]
 RP PROTEIN SEQUENCE OF 1-24.
 RC TISSUE=T-cell;
 RX MEDLINE=21482004; PubMed=11598101;
 RX DOI=10.1128/JAI.69.11.7190-7193.2001;
 RA Lee S.-U., Ferens W., Davis W.C., Hamilton M.J., Park Y.-H., Fox L.K.,
 RA Nagasens J., Bohach G.A.;
 RT "Identity of activation molecule 3 on superantigen-stimulated bovine
 RT cells is CD26.";
 RL Infect. Immun. 69:7190-7193(2001).
 RN [4]
 RP PROTEIN SEQUENCE OF 537-546.
 RC TISSUE=Kidney;
 RX MEDLINE=98293306; PubMed=9629661; DOI=10.1016/S0305-0491(97)00327-1;
 RA Ben-Shoshan I., Parola A.H.;
 RT "The CP-I subunit of adenosine deaminase complexing protein from calf
 RT kidney is identical to human, mouse, and rat dipeptidyl peptidase
 RT IV.";
 RL Comp. Biochem. Physiol. 119B:289-292(1998).
 CC -1- FUNCTION: Removes N-terminal dipeptides sequentially from
 CC polypeptides having unsubstituted N-termini provided that the
 CC penultimate residue is proline. Binds and regulates the activity
 CC of ADA.
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
 CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
 CC Zaa is neither Pro nor hydroxyproline.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
 CC soluble form (By similarity).
 CC -1- TISSUE SPECIFICITY: Intestinal epithelium, dendritic cells and
 CC several immune system tissues.
 CC -1- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
 CC by proteolytic processing (By similarity).
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AF461806; AAL67836.1; -; mRNA.
 CC EMBL; AY056834; AAL23628.1; -; mRNA.
 CC HSSP; P27487; 1PFO.
 CC SMR; P81425; 38-764.
 CC MEROPS; S09.003; -.
 CC InterPro; IPR002471; Pept_S9_AS.
 CC InterPro; IPR001375; Peptidase_S9.
 CC InterPro; IPR002469; Peptidase_S9B.
 CC InterPro; IPR000379; Ser_estra.
 CC Pfam; PF00930; DPPIV_N; 1.
 CC Pfam; PF00326; Peptidase_S9; 1.
 CC PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 CC AminoPeptidase; Direct protein sequencing; Glycoprotein; Hydrolase;
 KW Protease; Serine protease; Signal-anchor; Transmembrane.
 FT CHAIN 1 765
 FT FT Dipeptidyl peptidase 4 membrane form.
 FT FT Dipeptidyl peptidase 4 soluble form (By
 FT FT similarity).
 FT FT Cytoplasmic (Potential).
 FT FT Signal-anchor for type II membrane
 FT FT protein (Potential).
 FT FT Extracellular (Potential).
 FT FT Charge relay system (By similarity).
 FT FT ACT_SITE 629 629
 FT FT Charge relay system (By similarity).
 FT FT ACT_SITE 707 707
 FT FT Charge relay system (By similarity).
 FT FT ACT_SITE 739 739
 FT FT Charge relay system (By similarity).
 FT FT N-linked (GlcNAc...) (By similarity).
 FT FT CARBOHYD 84 84
 FT FT N-linked (GlcNAc...) (By similarity).
 FT FT CARBOHYD 91 91

FT CARBOHYD 149 149 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 218 218 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 228 228 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 271 271 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 280 280 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 320 320 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 392 392 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 495 495 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 684 684 N-linked (GlcNAc...) (By similarity).
 FT DISULFID 384 393 By similarity.
 FT DISULFID 443 446 By similarity.
 FT DISULFID 453 471 By similarity.
 FT DISULFID 648 761 By similarity.
 SQ SEQUENCE 765 AA; 88369 MW; E32165421P43E116 CRC64;
 Query Match 89.2%; Score 3585; DB 1; Length 765;
 Best Local Similarity 89.3%; Pred. No. 2.6e-235;
 Matches 649; Conservative 41; Mismatches 37; Indels 0; Gaps 0;
 QY 13 SRKTYTLTDYLNKTYRLKLSRWISDHEVLYKQENNLVFNAEVGNSSVPLENSTPDEF 72
 DB 38 SRRTYTLADYLNKTYFRMKFYNLRVWSDEHYLYKQENNLVFNAEVGNSSVPLENSTPDEF 97
 QY 73 GHSINDYSIGPDGQFILLENNYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVWS 132
 DB 98 GHSINDYSVSPDRQYILFEYNYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVWS 157
 QY 133 PVGHKLAVVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVFAYSALWWSWP 192
 DB 158 SVGHKLAVVWNNDIYVKNEPNSQRIITWTGKDDVIYNGITDWTVEEVFAYSALWWSWP 217
 QY 193 NGTFLAYAQFNDTEVPLIEYSFVSDLSIQPKTVRPVPPKAGAVNPVKFVWNTDLSLS 252
 DB 218 NSTFLAYAQFNDTEVPLIEYSFVSDLSIQPKTVKIPYKAGAVNPVKFVWNTDLSLS 277
 QY 253 VTNATSIQITAPASMLIGDHYLCDVTWATERISLOWLRRIQNTSYVMDICDYDSSGRWN 312
 DB 278 NINATSQIIVPPGSLVIGDHYLCDVTWATERISLOWLRRIQNTSYVMDICDYDSSGRWI 337
 QY 313 CLVARQHIEMSTTCWGRFRPSPHPPTLDGNSFKIISNEEGYRHCYFQIDKDCCTPIT 372
 DB 338 SSVGRQHIETSTTGWGRFRPSPHPPTSDGNSFYKIIISNEEGYKHICHFQTDKNCCTPIT 397
 QY 373 KGTWEVIGIEALTSYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 432
 DB 398 KGAEVIGIEALTSYLYISNEYKMGPGGRNLYKIQLNDYTKVTCLSCELNPERCOYYS 457
 QY 433 VSPSKEAKYYQLRCGPGPLPYTLHSSVNDKGLRVLEDNSALDKQLQNVQMPKSLDPII 492
 DB 458 VSPSQEAKYYQLRCGPGPLPYTLHSSVNDKGLRVLEDNSALDKQLQNVQMPKSLDPIH 517
 QY 493 LNETKFWYQIMLPHPDKSKYPLLLDVYAGPCSQKADTVFRLNWTATLASTENIIVASP 552
 DB 518 LHGTKFWYQIMLPHPDKSKYPLLELVYAGPCSQKADAI FRLNWTATLASTENIIVASF 577
 QY 553 DGRSGYQGQDKIMHAINRRIGTTEVEDQIRAAQFQSKMGFVDNKR IAIWGSYSGYVTSM 612
 DB 578 DGRSGYQGQDKIMHAINRRIGTTEVEDQIRAAQFQSKMGFVDNKR IAIWGSYSGYVTSM 637
 QY 613 VLGSQGVFKCGIAGVAPVSRWEYDYSVYTERYMGLEPTPEDNLDHYRNSVTMSRAENFKQV 672
 DB 638 VLGSQGVFKCGIAGVAPVSKWEYDYSVYTERYMGLEPTPEDNLDHYRNSVTMSRAENFKQV 697
 QY 673 EYLLIHGTADDNVHFQOQSAISKALVDVGVDFQAWTTTDEDHGIIASSTAHOHIIYTHMSHF 732
 DB 698 EYLLIHGTADDNVHFQOQSAISKALVDAGVDFQSMWYTTDEDHGIIASSTAHOHIIYTHMSHF 757
 QY 733 IKOCFSL 739
 DB 758 LKOCFSL 764
 RESULT 5

DPP4_FELCA
ID DPP4_FELCA STANDARD; PRT; 765 AA.
AC Q9NZI7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP IV) (T-cell activation antigen CD26) (Contains: dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)).
DE Name=DPP4; Synonyms=CD26;
GN Felis silvestris catus (Cat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae; Felinae; Felis.
OC NCBI_TaxID=9685;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Periphereal blood;
RX MEDLINE=20094000; PubMed=10630304; DOI=10.1007/s002510050616;
RA Nishimura Y., Miyazawa T., Ikeda Y., Izumiya Y., Nakamura K., Sato E., Mikami T., Takahashi E.;
RT "Molecular cloning and sequencing of a cDNA encoding the feline T-cell activation antigen CD26 homologue.";
RL Immunogenetics 50:366-368(1999).
CC -|- FUNCTION: Removes N-terminal dipeptides sequentially from polypeptides having unsubstituted N-termini provided that the penultimate residue is proline (By similarity).
CC -|- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided Zaa is neither Pro nor hydroxyproline.
CC -|- SUBUNIT: Homodimer (By similarity).
CC -|- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a soluble form (By similarity).
CC -|- PTM: The soluble form (SDPP) derives from the membrane form (MDPP) by proteolytic processing (By similarity).
CC -|- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC EMBL; AB023952; BAA92344.1; -; mRNA.
CC HSP; P27487; IPFO.
CC SMR; Q9NZI7; 38-765.
CC MEROPS; S09.003; -.
CC InterPro; IPR002471; Pept S9 AS.
CC InterPro; IPR001375; Peptidase_S9.
CC InterPro; IPR002469; Peptidase_S9B.
CC InterPro; IPR000379; Ser_estr.
CC Pfam; PF00930; DPPIV_N; 1.
CC PROSITE; PS00708; PRO_ENDOPEP_SBR; 1.
KW Aminopeptidase; Glycoprotein; Hydrolase; Protease; Serine protease;
KW Signal-anchor; Transmembrane.
FT CHAIN 1 765 Dipeptidyl peptidase 4 membrane form.
FT CHAIN 38 765 Dipeptidyl peptidase 4 soluble form (By similarity).
FT TOPO_DOM 1 6 Cytoplasmic (Potential).
FT TRANSMEM 7 29 Signal-anchor for type II membrane protein (Potential).
FT TOPO_DOM 30 765 Extracellular (Potential).
FT ACT_SITE 629 629 Charge relay system (By similarity).
FT ACT_SITE 707 707 Charge relay system (By similarity).
FT ACT_SITE 739 739 Charge relay system (By similarity).
FT CARBOHYD 84 84 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 91 91 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 149 149 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 178 178 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 228 228 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 280 280 N-linked (GlcNAc...) (By similarity).

FT CARBOHYD 320 320 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 330 330 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 331 331 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 519 519 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 684 684 N-linked (GlcNAc...) (By similarity).
FT DISULFID 384 393 By similarity.
FT DISULFID 443 446 By similarity.
FT DISULFID 453 471 By similarity.
FT DISULFID 648 761 By similarity.
SQ SEQUENCE 765 AA; 88213 MW; 38FCE98A22B175D9 CRC64;
Query Match 88.6%; Score 3560; DB 1; Length 765;
Best Local Similarity 88.0%; Pred. No. 1.3e-233;
Matches 641; Conservative 47; Mismatches 40; Indels 0; Gaps 0;
QY 13 SRKTYTLTYLKYVRLKLYSLRWISDHELYLYKQENNLVFNAAEYVNSGVLENSTFDEF 72
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 97
QY 73 GHSINDYISIPDGFILLEYVVKQWRHSYASDYIDLNKKQLITEERIPNNTQWVTWS 132
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 157
QY 133 FVGHKLAYVMNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEVFSYALSWMSP 192
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 217
QY 193 NGTFILAYAQFNDTEVPLIEYFYSDESLOYKTVRPYKAGAVNPTVKFFVNTDLS 252
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 277
QY 218 KGTFLAYAQFNDTQVPLIEYFYSDESLOYKTVRPYKAGAVNPTVKLVFKTDNLNP 312
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 337
QY 253 VTNATSIOTAPASMLIGDHYLDCVTWATOEIRISLQWLRRIQNSYVMDICDYDESSGRWN 372
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 397
QY 313 CLVARQHIEMSTGWVGRFRPSEPHFTLDGNSFYKIIISNEGYRHI CYPQIDKKDCTFTT 432
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 457
QY 373 KGTWEVIGIEALTSYLYYISNEYKMGPGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 492
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 517
QY 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGCSQKADTVFRLNWTYLASTENIIVASF 552
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 577
QY 518 LNETKFWYQMLPPHFDTSKKYPLLLDVYAGCSQKADAI FRLNWTYLASTENIIVASF 612
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 637
QY 553 DGRSGYOGDKIMHAINRLCTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSM 672
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 697
QY 578 DGRSGYOGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDDKRIATWGSYGGVVTSM 732
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 757
QY 613 VLGSYGKFGKCIAGVAPVSRWEYDSVYTERTWGLPTPEDNLDHYRNSVTMSRAENFKQV 772
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 797
QY 638 VLGSYGKFGKCIAGVAPVSRWEYDSVYTERTWGLPTPDQNDLYYKNSVTMSRAENFKQV 822
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 847
QY 673 EYLLIHGTADDNVHFQQAISKALVDVGDFQAMWYTDDEHGIIASSTAHOIYTHMSHF 902
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 927
QY 698 EYLLIHGTADDNVHFQQAISKALVDGAGVDFQAMWYTDDEHGIIASGAPAOHIYTHMSHF 952
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 977
QY 733 IKQCFSLP 740
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 765
QY 758 IKQCFSLP 765
RESULT 6
DPP4_FIG
ID DPP4_FIG STANDARD; PRT; 766 AA.
AC P22411; Q866G2;
DT 01-AUG-1991 (Rel. 19, Created)

MEROB; S09_003; --
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estr_S9.
DR Pfam; PF00930; DppIV_N_I.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SRR; 1.
KW 3D-structure; Aminopeptidase; Direct protein sequencing; Glycoprotein;
KW Hydrolase; Protease; Serine protease; Signal-anchor; Transmembrane.
FT CHAIN 1 766 Dipeptidyl peptidase 4 membrane form.
FT FT 38 766 Dipeptidyl peptidase 4 soluble form.
FT FT 1 6 Cytoplasmic (Potential).
FT TOPO_DOM 7 27 Signal-anchor for type II membrane
FT TRANSHEM 7 27 protein (Potential).
FT FT 28 766 Extracellular (Potential).
FT ACT_SITE 630 630 Charge relay system (By similarity).
FT ACT_SITE 708 708 Charge relay system (By similarity).
FT ACT_SITE 740 740 Charge relay system (By similarity).
FT CARBOHYD 85 85 N-linked (GLNAC. .).
FT CARBOHYD 92 92 N-linked (GLNAC. .).
FT CARBOHYD 150 150 N-linked (GLNAC. .) (By similarity).
FT CARBOHYD 179 179 N-linked (GLNAC. .) (Potential).
FT CARBOHYD 219 219 N-linked (GLNAC. .) (By similarity).
FT CARBOHYD 229 229 N-linked (GLNAC. .).
FT CARBOHYD 279 279 N-linked (GLNAC. .).
FT CARBOHYD 321 321 N-linked (GLNAC. .).
FT CARBOHYD 393 393 N-linked (GLNAC. .).
FT CARBOHYD 685 685 N-linked (GLNAC. .) (Potential).
FT DISULFID 385 394 N-linked (GLNAC. .).
FT DISULFID 444 447 Disulfide bond.
FT DISULFID 454 472 Disulfide bond.
FT DISULFID 649 762 Disulfide bond.
FT CONFLICT 32 32 Missing (in Ref. 2).
FT SEQUENCE 766 AA; 88242 MW; 8800D520BAEA856D CRC64;
Query Match 88.3%; Score 3548; DB 1; Length 766;
Best Local Similarity 88.2%; Pred. No. 8.6e-233;
Matches 642; Conservative 46; Mismatches 40; Indels 0; Gaps 0;
QY 13 SRKTYTLDTLYKNTVRLKLYSLRWISDHVELYKQENNILVPNAEYGNSSVLENSTDEP 72
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
39 SRKTYTLDTLYKSTFRVKFYTLQWISDHELYKQENNILFPNAEYGNSIFLENSTDFEL 98
QY 73 GHSINDYSISPDGQFILLEYVYKWRHSTASYDIYDLNKRQLITEERIPNNQTQWTWS 132
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
99 GYSTNDYSVSPRQFTLFEEVYVYKWRHSTASYDIYDLNKRQLITEERIPNNQTQWTWS 158
QY 133 PVGHKLAYVNNDIYVKIBPNLPSYITWTGKEDIINGITDWYEEVFSAVSALMWSP 192
DB 159 PVGHKLAYVNNDIYVKNEPNLSSQRITWTGENVIYNGVTDWYEEVFSAVSALMWSP 218
QY 193 NGTFFLAQAQNDTEVLBIYSFYSDESLSQPKTVRYPYKAGAVNPVKPFVNTDSLSS 252
DB 219 NGTFFLAQAQNDTEVLBIYSFYSDESLSQPKTVRYPYKAGAENPTVKPFVDTRTLP 278
QY 253 VTNATSIQITAPASMLIGHYLCDVMTAQERISLOWLRRIQNVSYMDICDYDESSGRWN 312
DB 279 NASVTSYQIVPPASVLIGHYLCGVTWVTEERISLOWIRRAQNYIIDICYDESTGRWI 338
QY 313 CLVARQHIEMTGTGWGRFRPBPHFTLDGNSFYKIISNEEGYRHICYFOIDKKDCTFIT 372
DB 339 SSVARQHIEISTGTGWGRFRPAEPHTSDGNSFYKIISNEEGYKHICHFTQDKSNCTFIT 398
QY 373 KGTWEVIGIEALTSDYLYIISNEYKMGPGGRNLKYIQLSDYTQVTCLSCEINPERCOYYS 432
DB 399 KGAEVIGIEALTSDYLYIISNEHKMGPGGRNLRYIQLANDYTKVTCLSCINPERCOYYS 458
QY 433 VSFSKEAKYVQLRCSPGLPLXYTLTHSSVNDKGILRVLEDSALDKMLQNVMPSKKCLDPFI 492
DB 459 ASFSNNKAKYIQLRCFGFGGLPLYLTLLHSSSSSDKLRVLEDSALDKMLQDVMPSKGLDVIN 518
QY 493 LNETRFWYQMILLPPHPDKSKCYPLLLDVDYVAGPCSQKADTVFRLNWATYLASTENIIIVASF 552

```

Db      519 LHGTFKFWYQMLPPHFDKSKKYPVLLIEVYAGSCQKQVDVFRLSWATYLASTENIIVASF 578
Qy      553 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVGNKRIAIWGSYGGYVTSM 612
Db      579 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEATFRQFSKMGFVDDKRIAIWGSYGGYVTSM 638
Qy      613 VLGGSGVPCGKTAVAPVRWEYDSVYTERYMGLETPEDNLDHYRNSVTMSRAENFKQV 672
Db      639 VLGGSGVPCGKTAVAPVRWEYDSVYTERYMGLETPEDNLDHYRNSVTMSRAENFKQV 698
Qy      673 EYLLHGTADDNVHVFQSQAKISKALVDGVDFQAMWYTDDEDHGIASSTAQHLYTHMSHF 732
Db      699 EYLLHGTADDNVHVFQSQAKISKALVDAGVDFTQTMWYTDDEDHGIASNMAHQHLYTHMSHF 758
Qy      733 IKQCFSLP 740
Db      759 LKQCFSLP 766

RESULT 7
DPP4 RAT
ID      DPP4 RAT      STANDARD;      PRT;      767 AA.
AC      P14740;
DT      01-APR-1990 (Rel. 14, Created)
DT      01-FEB-2005 (Rel. 46, Last sequence update)
DE      Dipeptidyl peptidase 4 (SC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE      IV) (T-cell activation antigen CD26) (GP110 glycoprotein) (Bile
DE      canaliculus domain-specific membrane glycoprotein) (Contains:
DE      Dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane
DE      form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV
DE      soluble form); Dipeptidyl peptidase 4 60 kDa soluble form (Dipeptidyl
DE      peptidase IV 60 kDa soluble form)).
GN      Name=Dpp4; Synonyms=Cd26;
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC      Muroidae; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RX      MEDLINE=89123496; PubMed=2563382;
RA      Ogata S., Misumi Y., Ikehara Y.;
RT      "Primary structure of rat liver dipeptidyl peptidase IV deduced from
RT      its cDNA and identification of the NH2-terminal signal sequence as the
RT      membrane-anchoring domain.";
RL      J. Biol. Chem. 264:3596-3601(1989).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=88068516; PubMed=3479775;
RA      Hong W., Doyle D.;
RT      "cDNA cloning for a bile canaliculus domain-specific membrane
RT      glycoprotein of rat hepatocytes.";
RL      Proc. Natl. Acad. Sci. U.S.A. 84:7962-7966(1987).
RN      [3]
RP      NUCLEOTIDE SEQUENCE OF 1-40.
RX      MEDLINE=89034185; PubMed=3182821;
RA      Hong W.J., Doyle D.;
RT      "Membrane orientation of rat gp110 as studied by in vitro
RT      translation.";
RL      J. Biol. Chem. 263:16892-16898(1988).
RN      [4]
RP      PROTEIN SEQUENCE OF 28-58, AND TISSUE SPECIFICITY.
RX      MEDLINE=90228896; PubMed=1970322;
RA      McCaughan G.W., Wicks J.E., Creswick P.F., Gorrell M.D.;
RT      "Identification of the bile canaliculus cell surface molecule GP110 as
RT      the ectopeptidase dipeptidyl peptidase IV: an analysis by tissue
RT      distribution, purification and N-terminal amino acid sequence.";
RL      Hepatology 11:534-544(1990).
RN      [5]
RP      PROTEIN SEQUENCE OF 281-302, AND MUTAGENESIS OF GLY-629; TRP-630;
RP      SER-631; TRY-632 AND GLY-633.

```

```

RC      TISSUE=Kidney;
RX      MEDLINE=94128239; PubMed=7905271;
RA      Iwaki-Egawa S., Watanabe Y., Fujimoto Y.;
RT      "N-terminal amino acid sequence of the 60-kDa protein of rat kidney
RT      dipeptidyl peptidase IV.";
RL      Biol. Chem. Hoppe-Seyler 374:973-975(1993).
RN      [6]
RP      PROTEIN SEQUENCE OF 624-648.
RX      MEDLINE=92190188; PubMed=1347701;
RA      Ogata S., Misumi Y., Tsuji E., Takami N., Oda K., Ikehara Y.;
RT      "Identification of the active site residues in dipeptidyl peptidase IV
RT      by affinity labeling and site-directed mutagenesis.";
RL      Biochemistry 31:2582-2587(1992).
RN      [7]
RP      SIGNAL-ANCHOR.
RX      MEDLINE=90338089; PubMed=1974258; DOI=10.1083/jcb.111.2.323;
RA      Hong W., Doyle D.;
RT      "Molecular dissection of the NH2-terminal signal/anchor sequence of
RT      rat dipeptidyl peptidase IV.";
RL      J. Cell Biol. 111:323-328(1990).
CC      -!- FUNCTION: Removes N-terminal dipeptides sequentially from
CC      polypeptides having unsubstituted N-termini provided that the
CC      penultimate residue is proline.
CC      -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC      Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC      Zaa is neither Pro nor hydroxyproline.
CC      -!- SUBUNIT: Homodimer.
CC      -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC      soluble form.
CC      -!- TISSUE SPECIFICITY: Expressed in bile ducts and other epithelial
CC      brush borders (small intestine, kidney, colon, pancreatic duct);
CC      acinar structures in salivary glands; endothelial structures and T
CC      cell areas in thymus, spleen and lymph node.
CC      -!- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC      by proteolytic processing.
CC      -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC
CC      EMBL; J04591; AAA41096.1; -; mRNA.
CC      EMBL; J02997; AAA41272.1; -; mRNA.
CC      PIR; A39914; A39914.
CC      HSP; P27487; 1PFQ.
CC      SMR; P14740; 37-766.
CC      MEROPS; S09.003; -.
CC      Ensembl; ENSRNOG0000030763; Rattus norvegicus.
CC      RGD; 2515; Dpp4.
CC      GO; GO:0005624; C:membrane fraction; IDA.
CC      GO; GO:0005625; C:soluble fraction; IDA.
CC      GO; GO:0004274; F:dipeptidyl-peptidase IV activity; TAS.
CC      GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
CC      InterPro; IPR002471; Pept_S9_AS
CC      InterPro; IPR001375; Peptidase_S9.
CC      InterPro; IPR002469; Peptidase_S9B.
CC      InterPro; IPR00379; Ser_eatrs.
CC      Pfam; PF00930; DPPIV_N; I.
CC      Pfam; PF00326; Peptidase_S9; 1.
CC      PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW      Aminopeptidase; Direct protein sequencing; Glycoprotein; Hydrolase;
KW      Protease; Serine protease; Signal-anchor; Transmembrane.
FT      CHAIN      1      767
FT      CHAIN      37      767
FT      CHAIN      281      767
FT      CHAIN      1      6
FT      TOPO_DOM      1      28
FT      TRANSMEM      7      28
FT      TOPO_DOM      29      767
FT      ACT_SITE      631      631

```


	EMBL; U12604; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12605; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12606; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12607; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12608; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12609; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12610; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12611; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12612; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12613; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12614; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12615; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12616; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12617; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12618; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; U12619; AAA82213.1; JOINED; Genomic DNA.	
DR	EMBL; AK085370; BAC39434.1; -, mRNA.	
DR	EMBL; BC022183; AAH22183.1; -, mRNA.	
DR	HSSP; P27487; LPQ.	
DR	SMR; P28843; 37-759.	
DR	MEROPE; S09.003; -.	
DR	Ensembl; ENSMUSG00000305000; Mus musculus.	
DR	MGI; MGI:94919; Dpp4.	
DR	GO; GO:0016021; C:integral to membrane; TAS.	
DR	GO; GO:0046581; C:intercellular canalliculus; IDA.	
DR	InterPro; IPR002471; Pept_S9_AS.	
DR	InterPro; IPR001375; Peptidase_S9.	
DR	InterPro; IPR002469; Peptidase_S9B.	
DR	InterPro; IPR000379; Ser_setra_-.	
DR	Pfam; PF00930; DPPIV_N; I.	
DR	Pfam; PF00326; Peptidase_S9; 1.	
DR	PROSITE; PS00708; PRO_ENDOPEP_SER; 1.	
KW	AminoPeptide; Direct protein sequencing; Glycoprotein; Hydrolase;	
KW	Protease; Serine protease; Signal-anchor; Transmembrane.	
FT	Dipeptidyl peptidase 4 membrane form.	
FT CHAIN	1 760	
FT FT CHAIN	37 760	Dipeptidyl peptidase 4 soluble form (By similarity).
FT		Cytoplasmic (Potential).
FT TOPO_DOM	1 6	Signal-anchor for type II membrane protein (Potential).
FT FT TRANSMEM	7 28	Extracellular (Potential).
FT		Cys-rich.
FT TOPO_DOM	29 760	Charge relay system (By similarity).
FT COMPBIAS	285 466	Charge relay system (By similarity).
FT ACT_SITE	624 624	Charge relay system (By similarity).
FT ACT_SITE	702 702	N-linked (GLNAC. .) (By similarity).
FT ACT_SITE	734 734	N-linked (GLNAC. .) (By similarity).
FT CARBOHYD	83 83	N-linked (GLNAC. .) (Potential).
FT CARBOHYD	90 90	N-linked (GLNAC. .) (Potential).
FT CARBOHYD	213 213	N-linked (GLNAC. .) (By similarity).
FT CARBOHYD	223 223	N-linked (GLNAC. .) (By similarity).
FT CARBOHYD	315 315	N-linked (GLNAC. .) (Potential).
FT CARBOHYD	328 328	N-linked (GLNAC. .) (Potential).
FT CARBOHYD	514 514	N-linked (GLNAC. .) (By similarity).
FT CARBOHYD	579 579	N-linked (GLNAC. .) (By similarity).
FT DISULFID	379 388	By similarity.
FT DISULFID	438 441	By similarity.
FT DISULFID	448 466	By similarity.
FT DISULFID	643 756	By similarity.
SQ SEQUENCE	760 AA; 87437 MW; A5F644B46E4A3DF8 CRC64;	
	Query Match 84.3%; Score 3390; DB 1; Length 760;	
	Best Local Similarity 84.7%; Pred. No. 4.9e-222;	
	Matches 616; Conservative 57; Mismatches 50; Indels 4; Gaps 1	
Qy	13 SRKTYTLDTLYLKNTRYLKLRLSRIVSDHEVLVYIKQENNILVFNAEYGNSSVFLENSTFDHF 72	
Db	37 SRRTYSLADYLKSTFRVKSYSLGWVSDFEVLVYIKQENILLLNHAHGNSIFLENSTFESP 96	
Qy	73 GHINDYISIPDGQFILLEYNYVQWRHSYTASYDYIDLNKRLQITEIRIPNNQTWTVWS 132	
Db	97 GY----HSVSPDLFLVLELYNYVQWRHSYTASTNIYDVNKRQLITEEKIPNNQTITWS 152	
Qy	133 PVGHKLAYVWNNDIYVKIEPLNPYSRYRTWTWKEDIINGLTDWVEEVFSYSAWWS 192	

OS Agkistrodon halys brevicaudus (Korean lamasa snake) (Gloydus halys
OS brevicaudus).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Gloydus.
OX NCBI_TaxId=259325;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Venom gland;
RA Ogawa Y., Yanoshita R., Murayama N., Higuchi S., Samejima Y.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB158225; BAD06333.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_setr.
DR Pfam; PF00930; DPPIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 751 AA; 86155 MW; DAF219B4FEE7629A CRC64;

Query Match 62.7%; Score 2519.5; DB 2; Length 751;
Best Local Similarity 62.4%; Pred. No. 9.2e-163;
Matches 454; Conservative 114; Mismatches 148; Indels 11; Gaps 5;

QY 14 RKTVTLDTYLVKNTVRLKYLRLWISDHELYLKQBNILVFNABYGNISVLENGSTFDFBG 73
DB 36 RRKESLEDYLSDEFQYKSYNLRWMSGHEVYVYTNQNVLLYNIDERSIVLSNLTLSFN 95

QY 74 HSINDYSISPDGQILLEYNVYKQHRHSYTSASYDIYDLNKRQLITEIRIPNNTQVWTSP 133
DB 96 SS--QAILSPDKFALLQYSYKVRHSYTSASYHYDLNNRKTITENPLPTNIQVSWSP 153

QY 134 VGHKLVVWNNDIYVKLEPNLPSYRITWTGKEDIYNGITDWYVEEVEVPSAYKSLWSPN 193
DB 154 VGHKLVVYRNVYVXATPNASPVQIITENGAENKILNGLADWVYEEEMFGTHSALWSPN 213

QY 194 GTPLAYAQFNDTEVPLLEYSPVDESLOYPKTVVPYKCAVAPNTVKFFVNTDLSVV 253
DB 214 GRFLAFAEINDETPVMWEYSFYSEDITLQYPTKIPIPKAGAINPTLRFLVIDI----SL 269

QY 254 TNATSIQITAPASMLIGDHYLCYDTWATQERISLQWLRRIQNYSYMDICYDESSGRWNC 313
DB 270 SPKNISEIVAPSSIISGDHYLSVVTWTDERICVQWLRRIQNFSVLTICDY--SGAWHC 326

QY 314 LVARQHLEMTSTGWGRPRPSEPFIPTLDGNSFYKIIISNEEGYRHCYFQIDKQCTFITK 373
DB 327 PKEREHLEESKTGWGRFPQSEPFTSDKISYVRIISDSEGYKHIHYTDSAGK-VKPIIS 385

QY 374 GTWEVIGTEALTSYLYIISNEYKMGDGRNLYKQLSDYTKVTCLSCLNPERCOYYSV 433
DB 386 GKWEVISISATNNLSYFISNEFGRPGRGRHLYKVDLNDKKITCITCNSEKCAQCFPSV 445

QY 434 SFSKEAKYQLRCSGPGPLLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSEKLLDFIL 493
DB 446 SFTSDSRYYKLCYGPDLPYFLQNSITDKAIKTLEDNNNLKNVLKEIQMPCKRLSNITL 505

QY 494 NETKFWYQMLPHFPDKSKYPLLLDVVYAGPCSKADTVFRLNWTATYLASTENIIVASFD 553
DB 506 HGQYTWYQMLPPNFDESCKYPLLIDVYVAGPCSKADAAFRINNSTVYLASSEIGIIVASFD 565

QY 554 GRGSGYQGDKTMHAINRPLGTFFVEDQIEAARQSKMGCFVDNKRITATWGSGYGVYTSMV 613
DB 566 GRGSGYQGDKTLHAIYRLRGTYEVEDQISAAKLFSEMSFVDKDRITATWGSGYGVYTSMV 625

QY 614 LGSGSGYFKCCGIAVAPSRWEYVDSVYTERYMGFLPTPEDNLDHYRNSVTVMGRAEFNKQVE 673

```
Db 273 MLPDFNSTEISPAIBKGDHYLSVVTWTERICLQWLRLRIQNTSVLTICDFESATGNW 332
Qy 312 NCLVARQH---IEMSTGWGRFRPSEPHFTLDGNSFYKIKISNEGYRHICVQFDKDC 368
Db 333 TC--HRKNFNKVKQLAGW--QISALCLPAPDNTTYKVSNTGYKHIIHYNGTEAPV 388
Qy 369 TFIKGTWEVIGIEALTSYLYISNEYKMGPGGRNLKYIQISDYTKVT--CLSCELNPER 427
Db 389 P-ITGKWEVISIAAVTKYFLYISNQNGEMPGGRNLKMLLESSEPKSTQCVCSDLNQR 447
Qy 428 COYTSVSSEKAKYQOLRCSPGLPYTLHLSVNDKGLRVLEDSALDKMLQNVOMPSKK 487
Db 448 COYTSASFSKQAYQOLNCLGPGMLHRSDDQVLRYLENNTELENSLKDIOQPSKK 507
Qy 488 LDFILNFKWYQMLPPHFDKSKYKPLLDVYAGPCSKQADTVRLNWNATYLASTENI 547
Db 508 LGSITVGGYNLYQMLPHLDSSKKYELLEVIYAGPCSKQDVHFRINWATYLASTEQI 567
Qy 548 IVASPDGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSGMFGVNDKRIAINWGSYGG 607
Db 568 IVASPDGRSGYQGDIEIMHAINRRLGTFEVEDQIEAARTFSEMSFVDKRIAINWGSYGG 627
Qy 608 YVTSNVLGSGVPCGKGIAPVSRWYDYDYTERYNGMLPTPEPNDLHYRNSVTMSRAE 667
Db 628 YVTSNVLGSGVPCGKGIAPVSRWYDYDYTERYNGMLPTPEPNDLHYRNSVTMSRAE 687
Qy 668 NFKOVEYLLIHTGADNVHFOQSAQISKALVDGVDFQAMWYTDSDHGIASSTAHOIYIT 727
Db 688 KFEVEYLLIHTGADNVHFOQSAQISKALVDAEVDFOAMWYTDKHGI--SQAHKIYIT 746
Qy 728 HNSHFIKQCFSLP 740
Db 747 HNSHFIKQCFSLP 759

RESULT 12
Q641D6_XENLA PRELIMINARY; PRT; 737 AA.
AC Q641D6;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE LOC397888 protein.
GN Name=LOC397888;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Uedini T.B., Toshiyuki S., Carninci P., Prange C.,
RA Reha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
```

```
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield J.S.N., Krzyzinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Klein S., Gerhard D.S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBAJ databases.
DR EMBL; BC082401; AAH82401.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR002469; Peptidase S9B.
DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00930; DPIP_N; 1.
DR Pfam; PF00326; Peptidase S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 737 AA; 84184 MW; 1E08CF94BDFBCC4B CRC64;

Query Match 58.4%; Score 2348; DB 2; Length 737;
Best Local Similarity 59.0%; Pred. No. 4.3e-151;
Matches 430; Conservative 111; Mismatches 168; Indels 20; Gaps 9;

Qy 14 RKTVTLTDYLNRYRLKLYSLRWISDHELYKQBNILVFNARYGNSVPLENSTFDFBG 73
Db 25 RKTFTLEDYFSDYRPRKSPGLKWSSENFVPRDKDNLVFNVDNETTETISNTTI---- 80

Qy 74 HSIND--YSISPDQFILLENNYKQWRHSYASDIYDLNKRQIITEERIPNNTQWTV 131
Db 81 HNSNSFYTLSEDRKALQYNEKLRHSYASDIYDIKKEIVAANLEPNKIQTW 140

Qy 132 SPVGHKLAYVWNDIYVKIRBNLPSYRIWTGKEDIYNGITDWWYREVEFSAYSAWWS 191
Db 141 SPVGHKLAYVWNNIYIKFVPGGISTITTINGENKILNGIPDWVYEEMFSTVALWWS 200

Qy 192 PNGTFLAYAFNDTEVPLIIBSYFSDLSQYPKTVRVPYPKAGAVNPTVKFVFNVTDSL 251
Db 201 PDATSLAYVERNDTDVPIEVSFYGEDSDQYFHTVPIPKAGARNPTVRLFAVNTSLA 260

Qy 252 SVTNATSIQTAPASMLIGHYLCVDTWATQERISLQWLRRIQNYVNDICDYDESSGRW 311
Db 261 VI---NPVEILPPEELSIDHYISGINWNTDKMAVQWLRRIQNVSLITMCE---GA 313

Qy 312 NCLVARQHIEWSTTGWGRFRPSEPHFTLDGNSFYKIKISNEGYRHICVQFDKDCFT 371
Db 314 NC---QPPVYSQSSGTGWGYPQSPAPYFDGLKTYIKLISNKGKYLHLFB-GSKDP 370

Qy 372 TKGTEVIGIEALTSYLYISNEYKMGPGGRNLKYIQL--SDYTKVTCLSCELPERCQ 430
Db 371 TSGNWEVTSIATVASNLFYVSNE--GFPGRRLYKIRLDSGYSYSAQCVCNTQRCCQ 428

Qy 431 YSVSFSKAKYQOLRCSPGLPYTLHLSVNDKGLRVLEDSALDKMLQNVOMPSKKLDF 490
Db 429 YSAFYSKNSKYSLNCGPGLPIYTVYVNSSNDETTRTMEDNEDLKLLEDEIOMPTKEN 488

Qy 491 IILNETKFWQMLPPHFDKSKYKPLLDVYAGPCSKQADTVRLNWNATYLASTENI 550
Db 489 IILDGFLWQLTLPHPHFDKSKYKPLLDVYAGPCSKQADTVRLNWNATYLASTE 548

Qy 551 SFDGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSGMFGVNDKRIAINWGSYGG 610
```

```

Db      549 SLDRGSGYQGGKIMHQIYHKLGTLEVDQDITAAKHFFSLGFVDPKMAINWSYGGYV 608
QY      611 SMVLGSGGVFKGIAVAPVSRWEYYDSYTRYMGLPTPEDNLDHYRNSTVMSRAENPK 670
Db      609 SMVLGSGSLFKGIAVAPVSMWHYYDSYTRYMGLPTPEDNLDNLSSTVMAQAQFK 668
QY      671 QVEYLLIHGTADDNVHFOQAISKALVDVGVDFQAMWYTDDEHGTASSTAHOHIYTHMS 730
Db      669 DVEYLLIHGTADDNVHFOQAANHSKALVDAQVDFTMWTYTDKDHGI--GGTANRHHYTHMS 727
QY      731 HFIKQCFSL 739
Db      728 HFLKQCFNI 736

RESULT 13
P70092_XENLA PRELIMINARY; PRT; 748 AA.
AC P70092;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Dipeptidyl-peptidase IV (EC 3.4.14.5).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
NCBI_TaxID=8355;
RN [1] NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=97390076; PubMed=9249015;
RA Vlaek R., Vilas U., Strobl B., Kreil G.;
RT "cDNA cloning and expression of secreted Xenopus laevis dipeptidyl
RT aminopeptidase IV.";
RL Eur. J. Biochem. 247:107-113(1997).
DR EMBL; Y08932; CAA70136.1; -, mRNA.
DR HSSP; P27487; IPFQ.
DR MEROPS; S09.003; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrb.
DR Pfam; PF00930; DPeIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 748 AA; 85587 MW; 31CCF61B25104E5B CRC64;

Query Match 58.0%; Score 2331.5; DB 2; Length 748;
Best Local Similarity 58.8%; Pred. No. 5.8e-150;
Matches 430; Conservative 112; Mismatches 164; Indels 25; Gaps 11;

QY 14 RKTTLTDYKNTYRLKLYSLRWISDHELYLKOENILVFNAEYGNSSVFLENSTDFEG 73
Db 37 RKTFLLEDYSDYRKPFGKLRWSENFVFRDKONVLFNVDNETTTIISNTTI---- 92

QY 74 HSGIND--YGISPDGQFLLILEYVVKQHRHSYASYDIYDLNKRQLITEIRPNNTQWTV 131
Db 93 HNSNSSFYTLSEDRKY-ALQYNYEKLWRHSYASYHYDIEKKEIVAANELPNKIQYITW 151

QY 132 SPVGHKLAVVWNNDIYVKEPNLPSYRITWTGKEDIYNGITDQVTEVEEYFSAYSALWWS 191
Db 152 SPVGHKLAVVWNNIYIKEVPGGISTITTTNGEHNKILNGIPDWYEEEMFSTNYALWWS 211

QY 192 PNGTFLAYAQFNDTEVPLIEYSFYSDESLSQYPKTVRPVYPKAGAVNPTVKFFVWNTDLS 251

```

```

Db      212 PDATSLAYFENDTDVPIEYSFYGEDSDQYPTVTVIYPKAGARNPTVRLFAVNTSLA 271
QY      252 SVTNATSIQTAPASMLIGDHYLCDVTWATQERISLQWLRRIQNTYSWMDICDYDESSGRW 311
Db      272 VI---NPVEILLPPEBELRSIDHYISGINWVTNKMVAQVLRRIQNTYSLTMCB----GAAM 324
QY      312 NCL--VARQHIEMSTGVVGRFRPSEPHFTLDGNSFYKISNEEGVYRHICYFQIDKKOCT 369
Db      335 NCQPPVYKQ----STGWGVFPQAPYFTQGLKYKLIISNKGYKXHLHFE--GSKDPV 379
QY      370 FITKGTWEVIGIEALTSYLYISNEYKMGPGGRNLYKIQL--SDYTKVTCLSCELNPERC 428
Db      380 AITSGNWEVTSIATVASNFLYVSNE--GPGRRQLYKIRLDSGYSACVTCNTRQERC 437
QY      429 QYYSVSFSKEAYIOLRCGPGCLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSKL 488
Db      438 QQYSAYFSKSKYSLNCGPGLPIYTVNSSNDNETRMTMEDNCLKLLDEIQMPTKEN 497
QY      489 DFIILNETKFWYOMILPPHEDSKKYPLLLDVYAGPCQKADTVFRLNWTATLASTENII 548
Db      498 KSIILIDGFEWYQLTLPPHFDKSKYPLLLIDVYGGPGSKVDQFFRLNWTATLASTEKII 557
QY      549 VASFDGRSGYQGGKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIHWGSYGY 608
Db      558 VASLDGRSGYQGGKIMHQIYHKLGTLEVDQDITAAKHFFSLGFVDPKMAINWSYGGY 617
QY      609 VTSMWLGSGGVFKGIAVAPVSRWEYYDSYTRYMGLPTPEDNLDHYRNSTVMSRAEN 668
Db      618 VTSMWLGSGSLFKGIAVAPVSMWHYYDSYTRYMGLPTPEDNLDNLSSTVMAQAQ 677
QY      669 FKQVEYLLIHGTADDNVHFOQAISKALVDVGVDFQAMWYTDDEHGTASSTAHOHIYTH 728
Db      678 FKQVEYLLIHGTADDNVHFOQAANHSKALVDAQVDFTMWTYTDKDHGI--GGTANRHHYTH 736
QY      729 MSHFIKQCFSL 739
Db      737 MSHFLKQCFNI 747

RESULT 14
Q53TP5_HUMAN PRELIMINARY; PRT; 760 AA.
AC Q53TP5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein FAP.
GN Name=FAP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
NCBI_TaxID=9606;
RN [1] NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RA Cotton M., Maupin R., Hawkins M., Harkins R.;
RT "The sequence of Homo sapiens BAC clone RP11-576I16.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2] NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3] NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [4] NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007750; AAY24205.1; -, Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.

```

DR	GO:0003824;	F: catalytic activity; IEA.		
DR	GO:0004274;	F: dipeptidyl-peptidase IV activity; IEA.		
DR	GO:0004287;	F: prollyl oligopeptidase activity; IEA.		
DR	GO:0006508;	P: proteolysis and peptidolysis; IEA.		
DR	InterPro: IPR002471;	Pept S9_AS		
DR	InterPro: IPR001375;	Peptidase_S9.		
DR	InterPro: IPR002469;	Peptidase_S9B.		
DR	InterPro: IPR000379;	Ser_estras.		
DR	Pfam: PF00930;	DPPIV N; I.		
DR	Pfam: PF00326;	Peptidase S9; 1.		
DR	PROSITE: PS00708;	PRO ENDOPEP SER; 1.		
KW	Hydrolase, Hypothetical protein, Transmembrane.			
SE	SEQUENCE 760 AA; 87713 MW; 7F817B5A4F75142 CRC64;			
Query Match				
Best Local Similarity 52.6%; Pred. No. 2.7e-139;				
Matches 390; Conservative 139; Mismatches 198; Indels 14; Gaps 7;				
Qy	3	PGSGHHHHSKRYTTLTDYLNKTVRLKLYSLRWISDHEHYLKO-ENNILVFNAYGNSS 61		
Db	28	PSRVHNSSENTMRALTLDKILNGTFSYKTPFPNMTSGOEYLFHQSDANNVILNYIETGQSY 87		
Qy	62	VFLNSNFDREGHIN--DYSISPDGQFLLLEYNVVKWRHSYTSYDIYLNKRLQITE 119		
Db	88	TILSNRTM----KSNVASNYGLSPDRQVYLESKLMRYSYTATYIYDLSNGEYVRG 143		
Qy	120	ERIPNNQWVTWSPGVGHKLAVVWNNDIYVKLEPNLPSVRIWTGKEDIYNGITDWWYEE 179		
Db	144	NELPRPIQYLCWSPVSKLAVYQNNIYVKORPGDPPQITFNGRENKIFNGIPDWWYEE 203		
Qy	180	EVFSAYSALMSPNGTFLAYAQFNDTEVPLIIEYGFYSDIESLQYPKTVRVPYPKGAVNPT 239		
Db	204	EMLATKVALMSPNGKFLAYAEFNDTDPVLAISYVGDE--QYPRITINIPYPKGAKNPV 261		
Qy	240	VKPFVNTDSLSSVTNATSIQITAPASMLIDGHYLCVDVWATQERISLOWLRRTQNTSYM 299		
Db	262	VRIPIDITTPAYYGPQ---EVPFAMTASSDYFWSLTTWTYDERVCLQWLKRVQNVSVL 318		
Qy	300	DICDYDESSGRNCLVARQHTEMSTGTWGVFRPSEPHFTLDGNSFYKIIISNEGYRHIC 359		
Db	319	SICDFREDQWDCPKTQEHIESERTGWAGGFFVSTVPFVSDAISYKIKFSDKQGYKHIH 378		
Qy	360	YFQIDKDKCTFITKGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQISDY-TKVTC 418		
Db	379	YIKDTVENAIQITSGKWEAINIPIVTDQLFYSNPEEYDGRNRYISIGSYPPSKKC 438		
Qy	419	LSCELNPERCOYYSVSKEAKYIQLRCSGPGCLPLYTLHSSVNDKGLRVLEDNSALDKML 478		
Db	439	VTCHLRKRCQYTTASFSDYAKYALVCYGGPISTLHDGRTDQEKILBENKELENAL 498		
Qy	479	QNVQPSKKLDIFILNETKFWYQMLPDPKSKYPLLLDVYAGPCSAKADTVFRLNWA 538		
Db	499	KNTQLPKEEIKLEVDHITLWYKMLPQDPRSKYPLLQVYGGPCSQSVRSFVFNWI 558		
Qy	539	TYLASENIIVASPDGRSGYQGDKIMHAINRRLGTPEVEDQIEBAARQFSKMGFVNDKRI 598		
Db	559	SYLASKEGVIALVDGRGTAFQGDKLLVAVYKLVGVIEVEDQITAVRKFIEMGFIDEKRI 618		
Qy	599	AIWGWSYGGYVTSWMLSGSGVFKGIAVAPSVWEYDYSVYTRYWGLPTTPBNLDHYR 658		
Db	619	AIWGSYGGYVSSILASGTLGPKGIAVAPVSSWEYYSVYTRFMRGLPTKDNLEHYK 678		
Qy	659	NSTVMSRAENPKQVEYLLIHGTADDNVHFQSAQISKALVDVGDVFOAMWYTTDSDHGIAS 718		
Db	679	NSTVMARAEYFRNVYDILLIHGTADDNVHFQNSAQIALVNAQVDFQAMWYSDQNGHL-S 737		
Qy	719	STAFQHIYTHMSHFIKQCFSL 739		
Db	738	GLSTNHLIYTHMFLKQCFSL 758		
RESULT 15				
SEPR HUMAN				

RESULT 15
SEPR HUMAN

RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]

RP PROTEIN SEQUENCE OF 192-208; 220-240 AND 510-521.

RX MEDLINE=94327249; PubMed=7519584;

RA Rettig W.J., Su S.L., Fortunato S.R., Scanlan M.J., Raj B.K.M.,

RA Garin-Chesa P., Healey J.H., Old L.J.

RT "Fibroblast activation protein: purification, epitope mapping and

RT induction by growth factors.";

RL Int. J. Cancer 58:385-392(1994).

CC -|- FUNCTION: May have a role in tissue remodeling during development
CC and wound healing, and may contribute to invasiveness in malignant
CC cancers.

CC -|- CATALYTIC ACTIVITY: Degrades gelatin and heat-denatured type I and
CC type IV collagen, but not native type I or type IV collagen. Does
CC not cleave laminin, fibronectin, fibrin or casein.

CC -|- SUBUNIT: Homodimer, or heterodimer with DPP4. The monomer is
CC inactive.

CC -|- SUBCELLULAR LOCATION: Type II membrane protein. Found in cell
CC surface lamellipodia, invadopodia and on shed vesicles.

CC -|- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1; Synonyms=L;

CC IsoId=Q12884-1; Sequence=Displayed;

CC Note=Major isoform;

CC Name=2; Synonyms=S, Truncated;

CC IsoId=Q12884-2; Sequence=VSP_005367;

CC -|- TISSUE SPECIFICITY: Fibroblast specific.

CC -|- INDUCTION: In fibroblasts at times and sites of tissue remodeling
CC during development, tissue repair, and carcinogenesis.

CC -|- PTM: N-glycosylated.

CC -|- PTM: The N-terminus may be blocked.

CC -|- SIMILARITY: Belongs to the peptidase S9B family.

CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----

DR EMBL; U09278; AAB49652.1; -; mRNA.

DR EMBL; U76833; AAC51668.1; -; mRNA.

DR EMBL; AF007822; AAF21600.1; -; mRNA.

DR EMBL; BC026250; AAH26250.1; -; mRNA.

DR PDB; 1z68; X-ray; A=39-757, B=-.

DR MEROPS; S09.007; -.

DR Ensembl; ENSG00000078098; Homo sapiens.

DR HGNC; HGNC:3590; PAP.

DR MIM; 600403; -.

DR GO; GO:0005887; C:integral to plasma membrane; TAS.

DR GO; GO:0030027; C:lamellipodium; IDA.

DR GO; GO:0005886; C:plasma membrane; NAS.

DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; NAS.

DR GO; GO:004222; F:metalloendopeptidase activity; TAS.

DR GO; GO:0042803; F:protein homodimerization activity; NAS.

DR InterPro; IPR002471; Pept_S9_AS.

DR InterPro; IPR001375; Peptidase_S9.

DR InterPro; IPR002469; Peptidase_S9B.

DR InterPro; IPR000379; Ser_estr.

DR Pfam; PF00930; DPPIV_N; I.

DR Pfam; PF00326; Peptidase_S9; 1.

DR PROSITE; PS00708; PRO_ENDOPSPSER; 1.

KW 3D-structure; Alternative splicing; Direct protein sequencing;

KW Glycoprotein; Hydrolase; Polymorphism; Protease; Serine protease;

KW Signal-anchor; Transmembrane.

FT TOPO_DOM 1 4 Cytoplasmic (Potential).

FT TRANSMEM 5 25 Signal-anchor for type II membrane

FT protein (Potential).

FT TOPO_DOM 26 760 Extracellular (Potential).

FT ACT_SITE 624 624 Charge relay system (By similarity).

FT ACT_SITE 702 702 Charge relay system (By similarity).

FT ACT_SITE 734 734 Charge relay system (By similarity).

FT CARBOHYD	49	49	N-linked (GlcNAc. . .)	(Potential).
FT CARBOHYD	92	92	N-linked (GlcNAc. . .)	(Potential).
FT CARBOHYD	99	99	N-linked (GlcNAc. . .)	(Potential).
FT CARBOHYD	314	314	N-linked (GlcNAc. . .)	(Potential).
FT CARBOHYD	679	679	N-linked (GlcNAc. . .)	(Potential).
FT VARSPLIC	1	521	Missing (in isoform 2).	
FT			/FTid=VSP_005367.	
FT	354	354	R -> T (in dSNP:1126507).	
FT			/FTid=VAR_019691.	
FT CONFLICT	207	207	P -> A (in Ref. 2 and 5).	
FT CONFLICT	229	229	K -> T (in Ref. 2 and 5).	
SQ SEQUENCE	760 AA;	87821 MW;	A0D34B4801BE07EA CRC64;	

Query Match 53.9%; Score 2168; DB 1; Length 760;
Best Local Similarity 52.5%; Pred. No. 8.1e-139;
Matches 389; Conservative 138; Mismatches 200; Indels 14; Gaps 7;

QY	3	PGSHHHHHHGRKTYTLDYLNKTVRLKYLRLWISDHEHYLYKQ-ENNLVFNABYGNSS	61
DB	28	PSRVHNSSEMTRALTLKDILNGTFSYKTFPPNNISGQEYLHQSDADNNIVLYNIETGQSY	87
QY	62	VLENSIFDEFGHSIN--DYSISPDGQFILLVYVKKWRHSYTSYDIYDLNKEQLITE	119
DB	88	TILSNRTW----KSVNASNYGLSPDRQFVLYESDKLWRYSYATYIYDLSNGFEVRG	143
QY	120	ERIPNNTOWTWSPVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWVYEE	179
DB	144	NELPRPIQYLCSWSPGVSKLAVYQNNIVLKQRPDGPFPQITFNGRENKIFNGIPDWVYEE	203
QY	180	EVFSAYSALWSPNGTFLAYAQFNDTEVPLIEYSYSDESLQYKTVRVPPKAGAVNPT	239
DB	204	EMLPTKYALWSPNGKFLAYAEFNDKDIPIVAYSYYGDE--QYPTINIPYPKAGAKNPV	261
QY	240	VKFFVWNTDSLSSVNTATSIQITAPASMLIGHYLCVDTWATQERISLQWLRRINYSVM	299
DB	262	VRIFIIDITYPAYVGPQ---EVPVPAMTASSDIYFSLTWVTVDERVCQWLKRVQNVSL	318
QY	300	DICDYDESSGWNCILVARQHIEMSTTGMVGRFPRSEPHFTLDGNSFYKIIISNEEGRHIC	359
DB	319	SICDFREDQWQDCPKTOEHIEESTGAGFFVSRPVFSYDAISYKIFSKDGYKHLH	378
QY	360	YFQIDKDCDCTFITKGTWEVIGIEALTSYLYISNEYKMGPGGRNLYKIQLSDY-TKVTC	418
DB	379	YIKDTVENAIQITSGWEAINIFRTQDSLFSYSSNEFEFEYFGRRNIYRISIGSYPPSKC	438
QY	419	LSCELNPERCOYYSVSFSKEAKYIOLRCGCLPLYTLHSSVNDKGLRVLEDNSALDKWL	478
DB	439	VTCHLRKERCOYITASFSADYAKYALVCYGGPIPISTLDGRTDQEIKILEENKELENAL	498
QY	479	QNVQMPSKKLDPIILNETKFWQMLPPHFKSKKYPLLLDYYAGPCSQKADTVFRLNWA	538
DB	499	KNIQLPKEIKKLEVDEITLWVKMLPPQFDRSKYPLLQYVGGPCSQSVRSVFAVNI	558
QY	539	TYLASTENIIVASFGRSGSGQDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKKI	598
DB	559	SYLASKEGWVIALVDGRGTAFOGDKLYAVYKRLGWVEVEDQITAVRKFIENGFIDEKRI	618
QY	599	AIWGSYGGYVTSWVLGSGGVFKGIAVAPVSRMEYVDSVYTERYMGILPTEDNLNHYR	658
DB	619	AIWGSYGGYVSSSLALASGTGLFKGIAVAPVSSMEYVASVYTERPMGLPTKDDNLEHYK	678
QY	659	NSTVMSRAENFKQVEYLHGTADDNVHVFQSAQISKALVDVGVDFQAMWYTDDEHGIAS	718
DB	679	NSTVWARAEYFENVLYLHGTADDNVHVFQSAQISKALVNAQVDFQAMWYSDQNHGL-S	737
QY	719	STAHOIYTHMSHFYKQCFSL	739
DB	738	GLSTNHLTYTHMTHFLKQCFSL	758

Search completed: February 15, 2006, 13:14:37

Job time : 237 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 13:10:58 ; Search time 44 Seconds
(without alignments)
1618.191 Million cell updates/sec

Title: US-10-659-055-3

Perfect score: 4020

Sequence: 1 ADPGSHHHHSKTYTLT.....AQHIYTHMSHFIRKQFSLP 740

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_80.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3933	97.8	766	1	CDHU26
2	3407.5	84.8	792	1	A39914
3	3386	84.2	760	1	S23752
4	1960.5	48.8	759	2	I38593
5	1129	28.1	803	2	I68600
6	1129	28.1	865	2	I54331
7	1106	27.5	803	2	A41793
8	950.5	23.6	793	2	T41703
9	914	22.7	818	1	A30107
10	799.5	19.9	711	2	S66261
11	745.5	18.5	829	2	J19514
12	733	18.2	931	2	A49737
13	725	18.0	799	2	T25174
14	720	17.9	779	2	T25173
15	628	15.6	738	2	A87516
16	594.5	14.8	741	2	JC5142
17	590.5	14.7	743	2	T37700
18	445.5	11.1	931	2	T32919
19	334.5	8.3	795	2	F82858
20	272	6.8	657	2	E70025
21	244.5	6.1	764	2	JC8016
22	226	5.6	622	2	F71174
23	221	5.5	591	2	H72474
24	219	5.4	642	2	C71137
25	214	5.3	632	2	H75057
26	205	5.1	631	2	E75007
27	205	5.1	683	2	E87495
28	191	4.8	709	2	B82580
29	188	4.7	659	2	F72568

ALIGNMENTS

RESULT 1

CDHU26

dipeptidyl-peptidase IV (EC 3.4.14.5) - human

N;Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence revision 23-Aug-1996 #text change 09-Jul-2004

C;Accession: S24313; B42408; A42408; B61136; S59510; I56154; S59857; S15520

R;Misumi, Y.; Hayashi, Y.; Arakawa, F.; Ikehara, Y.

R;Biochim. Biophys. Acta 1131, 333-336, 1992

A;Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a serine protease

A;Reference number: S24313; MUID:92329551; PMID:1352704

A;Accession: S24313

A;Molecule type: mRNA

A;Residues: 1-6, 'I', 8-766 <MIS>

A;Cross-references: UNIPROT:P27487; UNIPARC:UPI000016A880; EMBL:X60708; NID:G95335; PIR

R;Darmoul, D.; Lacasse, M.; Baricault, L.; Marguet, D.; Sapin, C.; Trotot, P.; Barbat, P.

J. Biol. Chem. 267, 4824-4833, 1992

A;Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cancer

IV mRNA levels during cell differentiation

A;Reference number: A42408; MUID:92165847; PMID:1347043

A;Accession: B42408

A;Molecule type: mRNA

A;Residues: 1-5, 'R', 7-436, 'S', 438-556, 'I', 558-662, 'E', 664-766 <DAR1>

A;Cross-references: UNIPARC:UPI000052ACB; GB:M60536; NID:G181569; PIDN:AAA52308.1; PIR

A;Experimental source: intestine

A;Note: this sequence corresponds with the author's translation

A;Accession: A42408

A;Molecule type: mRNA

A;Residues: 1-5, 'R', 7-436, 'S', 438-556, 'I', 558-662, 'E', 664-711, 'G', 713-766 <DAR2>

A;Cross-references: UNIPARC:UPI0000172A2B; GB:M60536; NID:G181569

A;Note: sequence extracted from NCBI backbone (NCBIN:83986, NCBI:83988); this sequence

R;Gorvel, J.P.; Ferrero, A.; Chambrault, L.; Rigal, A.; Bonicel, J.; Maroux, S.

Gastroenterology 101, 618-625, 1991

A;Title: Expression of sucrose-isomaltase and dipeptidylpeptidase IV in human small int

A;Reference number: A61136; MUID:91317403; PMID:1677636

A;Accession: B61136

A;Molecule type: protein

A;Residues: 1-15, 'X', 17-22 <GOR>

A;Cross-references: UNIPARC:UPI0000172A2C

R;Boehm, S.K.; Gum Jx., J.R.; Erickson, R.H.; Hicks, J.W.; Kim, Y.S.

Biochem. J. 311, 835-843, 1995

A;Title: Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation from a

A;Accession: S59510

A;Molecule type: DNA

A;Residues: 1-31 <BOE>

A;Cross-references: UNIPARC:UPI000016B4A6; GB:S79876; NID:G1195574; PIDN:AA835614.1; PIR

R;Tanaka, T.; Camerini, D.; Seed, B.; Torimoto, Y.; Dang, N.H.; Kameoka, J.; Dahlberg,

J. Immunol. 149, 481-486, 1992

A;Title: Cloning and functional expression of the T cell activation antigen CD26.

A;Reference number: I56154; MUID:92325476; PMID:1352530

A;Accession: I56154

A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-436,'S',438-766 <TAN>
A:Cross-references: UNIPARC:UPI000004F7BF; GB:M74777; NID:gl80082; PIDN:AAA51943.1; PID:
R:Abbott, C.A.; Baker, E.; Sutherland, G.R.; McCaughan, G.W.
Immunogenetics 40, 331-338, 1994
A>Title: Genomic organization, exact localization, and tissue expression of the human CD
A:Reference number: S59857; MUID:95012454; PMID:7927537
A:Accession: S59857
A:Molecule type: DNA
A:Residues: 1-436,'S',438-766 <ABB>
A:Cross-references: UNIPARC:UPI000004F7BF; EMBL:U13734
C:Genetics:
A:Gene: GDB:DP44
A:Cross-references: GDB:125239; OMIM:102720
A:Map position: 2q24.3-2q24.3
A:Introns: 2/3; 32/1; 65/1; 95/3; 122/3; 140/2; 164/3; 205/1; 258/3; 296/2; 341/3; 356/3
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidyl-peptidase hydrolase; glycoprotein; homodimer; proteinase; transmembr
F:1-6/Domain: intracellular #status predicted <INT>
F:7-28/Domain: transmembrane #status predicted <TAN>
F:29-766/Domain: extracellular #status predicted <EXT>
F:85,92,150,219,229,281,321,520,685/Binding site: carbohydrate (Asn) (covalent) #status
F:630,708,740/Active site: Ser, Asp, His #status predicted

Query Match	97.8%;	Score 3933;	DB 1;	Length 766;
Best Local Similarity	99.9%;	Pred. No. 6.5e-259;		
Matches 727;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY 13	SRKTYTLDTYKNTYRLKLYSLRWISDSHEYLKQENILVFNAYGNSVFLNSTFDEF	72		
Db	SRKTYTLDTYKNTYRLKLYSLRWISDSHEYLKQENILVFNAYGNSVFLNSTFDEF	98		
QY 73	GHSINDYSISPGQFILLBYNKKWRHSYTSYDIYDLNKKQLITEERIPNNTQVWTWS	132		
Db	GHSINDYSISPGQFILLBYNKKWRHSYTSYDIYDLNKKQLITEERIPNNTQVWTWS	158		
QY 133	PVGHKLAVYVNDIYVKIEPNLPSVRIITWTGKEDIYNGITDWWVEEVPFSAIWMKSP	192		
Db	PVGHKLAVYVNDIYVKIEPNLPSVRIITWTGKEDIYNGITDWWVEEVPFSAIWMKSP	218		
QY 193	NGTFLAYAQFNTEVPLIEYSFYSDESLOQPKTVRPYKAGVNPVTFVFNVDLSLS	252		
Db	NGTFLAYAQFNTEVPLIEYSFYSDESLOQPKTVRPYKAGVNPVTFVFNVDLSLS	278		
QY 253	VTNATSIQITAPASMLIGHYLCVDTWTATQERISQWLRRIQNYSVMDICDYDESSGRWN	312		
Db	VTNATSIQITAPASMLIGHYLCVDTWTATQERISQWLRRIQNYSVMDICDYDESSGRWN	338		
QY 313	CLVARKQHIEWSTGWGRPRPSEPHPTLDGNSFYKLIISNEEGVYRHICYFQIDKKCTFIT	372		
Db	CLVARKQHIEWSTGWGRPRPSEPHPTLDGNSFYKLIISNEEGVYRHICYFQIDKKCTFIT	398		
QY 373	KGTWEVIGTEALTSYLLIYSNEYKMGPCGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	432		
Db	KGTWEVIGTEALTSYLLIYSNEYKMGPCGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	458		
QY 433	VSFSEAKTYQIQRCSGPGPLPYTLTHSSVNDKGLRVLEDSALDKMLQNVQMSKKLDFII	492		
Db	VSFSEAKTYQIQRCSGPGPLPYTLTHSSVNDKGLRVLEDSALDKMLQNVQMSKKLDFII	518		
QY 493	LNETHFWYQMIPLPPHFDKSKYPLLLDLYVAGPCSKADTVFRNLNATYLASTENIIVASF	552		
Db	LNETHFWYQMIPLPPHFDKSKYPLLLDLYVAGPCSKADTVFRNLNATYLASTENIIVASF	578		
QY 553	DGRGSGYQGDKIMHAINRRLGTFEVEDQIEARQFSKMGFVNDKRIAIWGSYGGYVTSM	612		
Db	DGRGSGYQGDKIMHAINRRLGTFEVEDQIEARQFSKMGFVNDKRIAIWGSYGGYVTSM	638		
QY 613	VILGSGSGVPCGKIAPVAPSRWEYSDVYTERYVGLPTPDNDLDHYRNVSTVMSRAENFKQV	672		
Db	VILGSGSGVPCGKIAPVAPSRWEYSDVYTERYVGLPTPDNDLDHYRNVSTVMSRAENFKQV	698		

673 EYLLIHGTADDNVHFQSSAQISKALVDVGVDFQAMWYTDHGHGIASSTAHQHIYTHMSHF 732
|||||
699 EYLLIHGTADDNVHFQSSAQISKALVDVGVDFQAMWYTDHGHGIASSTAHQHIYTHMSHF 758
|||||
733 IKQCFSLP 740
|||||
759 IKQCFSLP 766
|||||

RESULT 2
A39914
dipeptidyl-peptidase IV (EC 3.4.14.5), membrane-bound form precursor - rat
N:Alternate names: GP110; membrane glycoprotein 110K; OX-61
N:Contains: dipeptidyl-peptidase IV, soluble form
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
A:Accession: A39914; A33315; A60730; A42203; S38949; A31781
R:Hong, W.; Doyle, D.
Proc. Natl. Acad. Sci. U.S.A. 84, 7962-7966, 1987
A>Title: cDNA cloning for a bile canalicular domain-specific membrane glycoprotein of ra
A:Reference number: A39914; MUID:88068516; PMID:3479775
A:Accession: A39914
A:Molecule type: mRNA
A:Residues: 1-792 <HON>
A:Cross-references: UNIPROT:P14740; UNIPARC:UPI000017098A; GB:J02997; NID:G204463; PIDN:
R:Ogata, S.; Miumi, Y.; Ikehara, Y.
J. Biol. Chem. 264, 3596-3601, 1989
A>Title: Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA an
A:Reference number: A33315; MUID:89123496; PMID:2563382
A:Accession: A33315
A:Molecule type: mRNA
A:Residues: 1-37,'A',39-182,'I',184-331,'T',333-351,'C',353-393,'V',395-561,'L',563-623,
A:Cross-references: UNIPARC:UPI0000129841; GB:J04591; NID:q203973; PIDN:AAA41096.1; PID:
A>Note: the authors translated the codon GCG for residue 38 as Arg, ACC for residue 332
A:Accession: B33315
A:Molecule type: protein
A:Residues: 1-20;35-54;427-443;505-509;511-520;530-538;593-600;602-608;618-627 <HO2>
A:Cross-references: UNIPARC:UPI0000172A37
A36; UNIPARC:UPI0000172A37
R:McCaughan, G.W.; Wikson, J.E.; Creswick, P.F.; Gorrell, M.D.
Hepatology 11, 534-544, 1990
A>Title: Identification of the bile canalicular cell surface molecule GP110 as the ectop
Quence.
A:Reference number: A60730; MUID:90228896; PMID:1970322
A:Accession: A60730
A:Molecule type: protein
A:Residues: 28-47,'XX',50-53,55-58 <MCC>
A:Cross-references: UNIPARC:UPI0000172A38
R:Ogata, S.; Miumi, Y.; Tsuji, E.; Takami, N.; Oda, K.; Ikehara, Y.
Biochemistry 31, 2582-2587, 1992
A>Title: Identification of the active site residues in dipeptidyl peptidase IV by affini
A:Reference number: A42203; MUID:94190188; PMID:1347701
A:Accession: A42203
A:Molecule type: protein
A:Residues: 'R',625-630,'X',632-648 <OG2>
A:Cross-references: UNIPARC:UPI0000172A39
R:Iwaki-Egawa, S.; Watanabe, Y.; Fujimoto, Y.
Biol. Chem. Hoppe-Seyler 374, 973-975, 1993
A>Title: N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl p
A:Reference number: S38949; MUID:94128239; PMID:7905271
A:Accession: S38949
A>Status: preliminary
A:Molecule type: protein
A:Residues: 281-302 <IWA>
A:Cross-references: UNIPARC:UPI0000172A3A
R:Hong, W.; Doyle, D.
J. Biol. Chem. 263, 16892-16898, 1988
A>Title: Membrane orientation of rat gp110 as studied by in vitro translation.
A:Reference number: A31781; MUID:89034185; PMID:3182821
A:Accession: A31781
A:Molecule type: mRNA
A:Residues: 1-40 <HO3>
A:Cross-references: UNIPARC:UPI0000172A3B

C:Comment: This protein is localized to the bile canalculus, which is the apical domain
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidyl-peptidase hydrolase; glycoprotein; homodimer; liver; serine protease
F:1-792/Product: dipeptidyl-peptidase, membrane-bound form #status experimental <MATW>
F:1-28/Domain: signal sequence #link MATS #status experimental <SIG>
F:1-6/Domain: intracellular #status predicted <INT>
F:7-28/Domain: transmembrane #status predicted <TMN>
F:29-792/Domain: extracellular #status predicted <EXT>
F:35-792/Product: dipeptidyl-peptidase, soluble form #status experimental <MATS>
F:83,90,148,217,227,319,521,686/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:631/Active site: Ser #status experimental
F:703,741/Active site: Asp, His #status predicted

Query Match 84.8%; Score 3407.5; DB 1; Length 792;
Best Local Similarity 84.5%; Pred. No. 3.2e-223;
Matches 617; Conservative 54; Mismatches 56; Indels 3; Gaps 1;

Qy	13	SRKTYTLTDYLVKNTYRLKLSLRWISDHEVLYKQENILVFNAYGVSSVPLENSTPDEF	72
Dd	37	SRRTYTLADYLVKNTFRVKSYSLRWVSDSEVLYKQENILVFNAYGVSSVPLENSTPDEF	96
Qy	73	GHSINDYSISPDGQFILLVYVYKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWTVS	132
Dd	97	GDSISDYSVSPDRFLFVLLVYVYKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWTVS	156
Qy	133	PVGHKLAVVWNNDIYVKIENPLPSYRIWTGKEDIYNGITDWYVEEVSFAYSALMWSP	192
Dd	157	QEGHKLAVVWNNDIYVKIENPLPSYRIWTGKEDIYNGITDWYVEEVSFAYSALMWSP	216
Qy	193	NGTFLAVAFQNDTVPLIIEYFYSDSLSQYPTKTVVPYKAGAVNPTKFFVNTDLSLS	252
Dd	217	NGTFLAVAFQNDTVPLIIEYFYSDSLSQYPTKTVVPYKAGAVNPTKFFVNTDLSLS	276
Qy	253	VTNATSIQITAPASMLGDHYLVDVWATQRIISLOWLRRIQNYSDMDICDYDESSGRWN	312
Dd	277	TTTTIPMQITAPASVTTGDHYLVDVWATQRIISLOWLRRIQNYSDMDICDYDESSGRWN	336
Qy	313	CLVARQHEIMSTGQVGRFRPSRPHFTLDGNSFYKISNERYGRIHICVFQIDKK---	369
Dd	337	CPTQHEIMSTGQVGRFRPSRPHFTLDGNSFYKISNERYGRIHICVFQIDKK---	396
Qy	370	FITKGTWEVIGIRALTSYLYISNEYKMPGGRNLYKIQISDYTKVTCISCLNPERCQ	429
Dd	397	FITKGTWEVIGIRALTSYLYISNEYKMPGGRNLYKIQISDYTKVTCISCLNPERCQ	456
Qy	430	YYSVSFSEAKYYQLRCSGPGPLPLYTLHSSVNDKGLVLENSALDKMLQVQMPKSLD	489
Dd	457	YYSVLSKBAKYYQLRCSGPGPLPLYTLHSSVNDKGLVLENSALDKMLQVQMPKSLD	516
Qy	490	FIILNETKFWYQMLPLPHFDKSKYPLLLDVYAGPCQKADTVFRLNWTYLASTENIIV	549
Dd	517	FIVLNETKFWYQMLPLPHFDKSKYPLLLDVYAGPCQKADTVFRLNWTYLASTENIIV	576
Qy	550	ASPDGRSGYGGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIATWGSYGGYV	609
Dd	577	ASPDGRSGYGGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIATWGSYGGYV	636
Qy	610	TSMWLGSQGVFKGCIAPVSRVEYVDSVYTERYMGILPTPEDNLDHYRNSVMSRAENF	669
Dd	637	TSMWLGSQGVFKGCIAPVSRVEYVDSVYTERYMGILPTPEDNLDHYRNSVMSRAENF	696
Qy	670	KQVEYLLHGTADDNVHFQQAQISKALVDGVDFQAWWYTTDHDGFIASSTAHQHIYTHM	729
Dd	697	KQVEYLLHGTADDNVHFQQAQISKALVDGVDFQAWWYTTDHDGFIASSTAHQHIYTHM	756
Qy	730	SHFIKQCFSL	739
Dd	757	SHFIKQCFSL	766

dipeptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse
N:Alternate names: CD26 alpha subunit; TRAM alpha subunit
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S23752; A46465; A56030
R:Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.
J. Biol. Chem. 267, 2200-2208, 1992
A:Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-
A:Reference number: S23752; MUID:92129288; PMID:1370813
A:Accession: S23752
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-760 <MAR>
R:Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.
J. Immunol. 147, 447-454, 1991
A:Title: Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptida-
A:Reference number: A46465; MUID:91302787; PMID:1712807
A:Accession: A46465
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <VIV>
A:Cross-references: UNIPARC:UPI0000172A2E
A:Experimental source: M14.T thymoma cells, Swiss nu/nu
A:Note: sequence extracted from NCBI backbone (NCBIP:42236)
R:Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.
Biochemistry 33, 15204-15214, 1994
A:Title: Structure of the mouse dipeptidyl peptidase IV (CD26) gene.
A:Reference number: A56030; MUID:95092780; PMID:7999781
A:Accession: A56030
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 746-760 <BER>
A:Cross-references: UNIPARC:UPI0000172A2F; GB:U12620
C:Genetics:
A:Gene: CD26
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidyl-peptidase hydrolase; glycoprotein; transmembrane protein
F:213,223,315,514,679/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:624,702,734/Active site: Ser, Asp, His #status predicted

Query Match 84.2%; Score 3386; DB 1; Length 760;
Best Local Similarity 84.6%; Pred. No. 8.6e-222;
Matches 615; Conservative 57; Mismatches 51; Indels 4; Gaps 1;

Qy	13	SRKTYTLTDYLVKNTYRLKLSLRWISDHEVLYKQENILVFNAYGVSSVPLENSTPDEF	72
Dd	37	SRRTYSLADYLVKNTFRVKSYSLRWVSDFEVLYKQENILVFNAYGVSSVPLENSTPDEF	96
Qy	73	GHSINDYSISPDGQFILLVYVYKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWTVS	132
Dd	97	GY-----HSVSPDRFLFVLLVYVYKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWTVS	152
Qy	133	PVGHKLAVVWNNDIYVKIENPLPSYRIWTGKEDIYNGITDWYVEEVSFAYSALMWSP	192
Dd	153	PEGHKLAVVWNNDIYVKIENPLPSYRIWTGKEDIYNGITDWYVEEVSFAYSALMWSP	212
Qy	193	NGTFLAVAFQNDTVPLIIEYFYSDSLSQYPTKTVVPYKAGAVNPTKFFVNTDLSLS	252
Dd	213	NNTFLAVAFQNDTVPLIIEYFYSDSLSQYPTKTVVPYKAGAVNPTKFFVNTDLSLS	272
Qy	253	VTNATSIQITAPASMLGDHYLVDVWATQRIISLOWLRRIQNYSDMDICDYDESSGRWN	312
Dd	273	SSNAPIQITAPASVARGDHYLVDVWATQRIISLOWLRRIQNYSDMDICDYDESSGRWN	332
Qy	313	CLVARQHEIMSTGQVGRFRPSRPHFTLDGNSFYKISNERYGRIHICVFQIDKKDCTFIT	372
Dd	333	CPSEQQHVENSTGQVGRFRPSRPHFTLDGNSFYKISNERYGRIHICVFQIDKKDCTFIT	392
Qy	373	KGTWEVIGIRALTSYLYISNEYKMPGGRNLYKIQISDYTKVTCISCLNPERCQYVS	432
Dd	393	KGAWEVISIRALTSYLYISNEYKMPGGRNLYKIQISDYTKVTCISCLNPERCQYVS	452

QY 433 VSFSEAKYQYLCRSGPLPLTLHSSVNDKGLRVLEEDNSALDKMLQNVQMPSSKKLDPLII 492
Db 453 VSFSEAKYQYLCRSGPLPLTLHSSVNDKGLRVLEEDNSALDKMLQNVQMPSSKKLDPLIV 512
QY 493 LNETFWYQMLPPHFDKSKKYPDLLDVYAGPCSKADTVFRLNWTYLASTENIIVASF 552
Db 513 LNETFWYQMLPPHFDKSKKYPDLLDVYAGPCSKADTVFRLNWTYLASTENIIVASF 572
QY 553 DGRSGYQGDKIMHAINRRLGTFTVEVDQIEAARQSKMGFVNDKRIAIWGWYGGYVTSM 612
Db 573 DGRSGYQGDKIMHAINRRLGTFTVEVDQIEAARQSKMGFVNDKRIAIWGWYGGYVTSM 632
QY 613 VLGSYGVPKCGIAPVPSRWYSDVYTERYMGLEPTREDNLHDYRNSVTMSRAENFKQV 672
Db 633 VLGSYGVPKCGIAPVPSRWYSDVYTERYMGLEPTREDNLHDYRNSVTMSRAENFKQV 692
QY 673 EYLLHGTADDNVHQQSAQISKALVDGVDFQAMWYTTDEDHGIIASSTAHOIYTHMSHF 732
Db 693 EYLLHGTADDNVHQQSAQISKALVDGVDFQAMWYTTDEDHGIIASSTAHOIYTHMSHF 752
QY 733 IKQCPSL 739
Db 753 LQCCPSL 759

RESULT 4
138593
fibroblast activation protein-alpha - human
N:Alternate names: FAP-alpha
C:Species: Homo sapiens (man)
C:Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 29-Aug-1997
C:Accession: I38593
R:Scanlan, M.J.; Raj, B.; Calvo, B.; Garin-Chesa, P.; Sanz-Moncasi, M.P.; Healey, J.; Ol
Proc. Natl. Acad. Sci. U.S.A. 91, 5657-5661, 1994
A:Title: Molecular cloning of fibroblast activation protein alpha, a member of the sexin
A:Reference number: I38593; MUID:94561645; PMID:7911242
A:Accession: I38593
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-759 <RES>
A:Cross-references: UNIPARC:UPI000175B7D; EMBL:U09278; NID:g507749; PID:g507750
C:Genetics:
A:Gene: GDB:FAP
A:Cross-references: GDB:374184; OMIM:600403
A:Map position: 2q23-2q23
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: transmembrane protein

Query Match 48.8%; Score 1960.5; DB 2; Length 759;
Best Local Similarity 48.5%; Pred. No. 4.8e-125; Indels 35; Gaps 11;
Matches 364; Conservative 138; Mismatches 214

QY 3 PGSSHHHHHSRTYTLTDYLNKTYRLKLYSLRWISDHELYLKQ-ENNILVFNAYEGNSS 61
Db 28 PSRVHSENTWRALTLDKLNCTFSYKTFPFWISGQYELHQSADNNVLVNIETGQSY 87
QY 62 VFLENSTDPGSHIN--DYSISPGQFILLEYNVYKQWRHSYASYDYLDLKNKQLITE 119
Db 88 TILSNRTM---KSVNASNYGLSPDRQFVYLSDSYSLKWRYSYATYTYDLSNGEFVRG 143
QY 120 BRIPNNTQVTPVGHKLYAVNNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWYEE 179
Db 144 NELPRIQYLCWSPVGSKLAYVQNNIYLKQRPDPPFQITFNGRENKIFNGIPDWYEE 203
QY 180 EYFSAYSALWSPNGTFLAYAFNDTEVPLIYSFYSDSLQYPKTVRPVYPKAGAVNPT 239
Db 204 EMLPTKVALWSPNGKFLAYAFENDKDIPIVAYSYYGDS--QYPTINIPYKAGAKNPV 261
QY 240 VKFFVNTDLSVNTATSIQITAPASMLIGHYLCDTVATQERISLOWLRRIQNYSYM 299
Db 262 VRIFIDITTPAYVGPQ---EVPVPAMIASSDYFSLWLTWTVDERYCLOWLKRQVNSYL 318
QY 300 DICVDYDESSGRNCLVARQHIEMSTGTGWRFRPSPHPTLDGNSPKYIISNEEGRHIC 359

Db 319 SICDFREDQWDCPKTQEHIEESRTGWAGFFVSRPVSFYDAISYKIFSDKDGXYKH 378
QY 360 YFQIDKDCDCTITKGTWVIGIEALTSYLYISNEYKMGPCGRNLYKIQLSDY-TKVT 418
Db 379 YIKDTVNAIQITSGKWEAINIFRVTDLSLFYSSNEFEYFGRRNIYRISIGSPPSKC 438
QY 419 LSCELNPERCOYVSFSKEAKYQYLCRSGPLPLTLHSSVNDKGLRVLEEDNSALDKML 478
Db 439 VTCHLRKERCQYITASPSDYAKYALVCYGPPISTLHDGRTDQEIILKEENKELENAL 498
QY 479 QNVQMPSSKKLDPLIILNETKFWYQMLPPHFDKSKKYPDLLDVYAGPCSKADTVFRLNWA 538
Db 499 KNIQLPKEIEKKLEVDKITLWYKMLPPQFDRSKKYPDLLQVYGPCSQSVRSVFAVNI 558
QY 539 TYLASTENIIVASFPGRSGYQGDKIMHAINRRLGTFTVEVDQIEAARQSKMGFVNDKRI 598
Db 559 SYLASKEGWIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRI 618
QY 599 AIWGSYG-GYVT-----SMVLGSGSGVFKCGIAPVPSRWYSDVYTERYMGLEP 648
Db 619 AIWGSYEIFRITGPCIWNWSFQMYSSGSL-QGILRVCLHR-----EIHGSP 667
QY 649 TPEDNLDRNSVTMSRAENFKQVEYLLIHGTADDNVHQQSAQISKALVDGVDFQAMW 708
Db 668 NKDDNLEHYKNSVTWARAEPFRVDYLLIHGTADDNVHQQSAQISKALVNAQVDQAMW 727
QY 709 YTDEHGTASSTAHOIYTHMSHFIKQCPSL 739
Db 728 YSDQNHGL-SGLSTNHLTYTHMTFLKQCPSL 757

RESULT 5
168600
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C:Accession: I68600
R:Yokotani, N.; Doi, K.; Wenthhold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-relat
A:Reference number: I54331; MUID:93372805; PMID:8103397
A:Accession: I68600
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-803 <RES>
A:Cross-references: UNIPARC:UPI000016A7F9; GB:M96860; NID:g306707; PIDN:AAA35761.1; PID:
C:Superfamily: dipeptidyl-peptidase IV

Query Match 28.1%; Score 1129; DB 2; Length 803;
Best Local Similarity 32.7%; Pred. No. 1.4e-68;
Matches 244; Conservative 155; Mismatches 303; Indels 44; Gaps 20;

QY 14 RRTYTLTDYLNKTYRLKLYSLRWISDHELYLKQ-ENNILVFNAYEGNSSVFLNSTDSEF 72
Db 66 KKKVTVEIDLPSDDPKIHDPKAWISDTEFIYREQGTGTVRLWNVETNTSTVLIEGKKIESL 125
QY 73 GHSINDYISPGQFILLEYNVYKQWRHSYASYDYLDL--NKRQLITEIRPN-NTQV 129
Db 126 -RAIR-YEISPDREYALFSYVPEYIYQHSYTGYYVLSKIPHGDPQSLDPEPSNAKLQYA 183
QY 130 TWSPVGHKLYAVNNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWYEEYFSAYSALW 189
Db 184 GWPGKQQLIFIFENNIYCAHVGHQKARVSTGKEGYINGLSLWLYEEELKTHIAHW 243
QY 190 WSPNGTFLAYAFNDTEVPLIYSFYSDSLQYPKTVRPVYPKAGAVNPTKFFVNTDS 249
Db 244 WSPDGTFLAYAAINDSRVPEIMELPTVTGSI--YPTVKPYHPKAGSENPSISLHVI---G 298
QY 250 LSSVTNATSIQITAPASMLIGHYLCDTVATQERISLOWLRRIQNYSYMDCVDYDESS 309
Db 299 LNPPTH--DLEMPMPDDPRMREYITVVKWATSTKVAVTWLNRAQNVSLTLC--DATTG 354

```
Qy 310 RNCLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEBGRHICYF----- 361
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
355 ----VCTKCH-EDSEAWLHR-QNEEPVFSKGRKFFPIRAIPQGRGKFFHITVSSSQP 408
Qy 362 QIDKDCFTFKGTWEVIGIRAL--TSDLYIYISNEYKMGPGGRNLKYIQLSDYTKVTCL 419
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
409 NSSNDNIQSIITSGDWDTKILAYDEKGNKIYFLSTE--DLPRRLQIYSANTEGNRQCL 466
Qy 420 SCENLPERCOYYSVFSKKAQYQLRCGPGPLPYTLHLSSVNDKGLRVLNDSALDKMLQ 479
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
467 SCDL- VENCYFSAFSHSDWDFLLKCEGPGVPMVTVHTTDDKKMPDETNEHVKKAIN 525
Qy 480 NVQMPKCLDPIILNETKFWOMILPPHFDKSKYPLLLDVYAGCSKADTVFRLNAT 539
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
526 DRQPKVEYRDIEIDDDNLPMLKLPATFTDTTHYPLLLVVDGTPGSGVAEKFEVSMET 585
Qy 540 YLASTENIIVASFGRGSGYQGDIMHAINRRLGTTFEVEDOIEAARQFSKMGFVNDKRIA 599
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
586 VMVSHGAVVVKCDGRSGFGQTKLLEHVRRLGLLEEKQMEAVRMLKEQYIDRTKVA 645
Qy 600 IWGWSYGGYVTSMLV---GSGSG-VFKCGIAPVPSRWEYDVSVYTRYMGLPTPEDNLD 655
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
646 VFGDYGYLSTYILPAKGENQGTFTCGSALSPTDFKLYASAFSEYRGL---HGLD 701
Qy 656 H--YRNSTVMGRARENFKQVEYLLIHGTADDNVHFPQQAISKALVDVGVDFQAMWYTD 713
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
702 NRAYEMTKVAHRVSALREQQFLIHPHTADEKIHFOHTAELITQLIRGKANYSLQIYPDES 761
Qy 714 HGIASSTAHQHIYTHMSHFIKQCPSL 739
Db 762 HYFTSSSLKQHLRYSIINFFVECPRI 787
```

RESULT 6

```
I54331
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I54331
R:Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A>Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related
A:Reference number: I54331; MUID:93372805; PMID:8103397
A:Accession: I54331
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-865 <RES>
A:Cross-references: UNIPROT:P42658; UNIPARC:UPI00000047378; GB:M96859; NID:g306705; PIDN:
C:Superfamily: dipeptidyl-peptidase IV
```

```
Query Match 28.1%; Score 1129; DB 2; Length 865;
Best Local Similarity 32.7%; Pred. No. 1.6e-68;
Matches 244; Conservative 155; Mismatches 303; Indels 44; Gaps 20;
```

```
Qy 14 RKTVTLTDYLNKTVRLKLYSLRWISDHYLYK-QENNILVFNARYGNSSVPLENSTPDEF 72
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
128 KKKVTVEDLFSDFDKIHPEAKWISDTEFYREQGTVRLNWNVTNSTVLIEGKKIESL 187
Qy 73 GHSINDYSISDGGFILLNYYVQWRHSYTSASDIYDL--NKRQLITEERIPN-NQOW 129
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
188 -RAIR-YEISPDREYALFSNVNVEPIYQHSYTGYYVLSKIPHGDPQSLDPPPEVSNAKLOYA 245
Qy 130 TWSVPVGHKLAVVWNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWWYEEVFSAYSALW 189
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
246 GWGPKGQQLIFIPENNIYYCAHVCKQAIRVVVSTGKEGVYINGLSDWLYEEBILKTHIAHW 305
Qy 190 WSPNGTFLAYAFNDTEVPLEIYSFYSDESIQLYPKTVVPYPKAGAVNPTVFFVNTDS 249
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
306 WSPDGTRLAVAAINDSRVPIWELPTYTSGI--YPTVTPYHPYKAGSENPISLHVI---G 360
Qy 250 LSSVTNATSIQITAPASMLIGDHYLCDVWATQBRISLOWLRRIQNTSVMDICDYDSSG 309
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
361 LNPETH--DLEMMPPDDPRMREYIITVWKATSTKVAVTMLNRAQNVSIILTLCDATG 416
```

```
Qy 310 RNCLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEBGRHICYF----- 361
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
417 ----VCTKCH-EDSEAWLHR-QNEEPVFSKGRKFFPIRAIPQGRGKFFHITVSSSQP 470
Qy 362 QIDKDCFTFKGTWEVIGIRAL--TSDLYIYISNEYKMGPGGRNLKYIQLSDYTKVTCL 419
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
471 NSSNDNIQSIITSGDWDTKILAYDEKGNKIYFLSTE--DLPRRLQIYSANTEGNRQCL 528
Qy 420 SCENLPERCOYYSVFSKKAQYQLRCGPGPLPYTLHLSSVNDKGLRVLNDSALDKMLQ 479
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
529 SCDL- VENCYFSAFSHSDWDFLLKCEGPGVPMVTVHTTDDKKMPDETNEHVKKAIN 587
Qy 480 NVQMPKCLDPIILNETKFWOMILPPHFDKSKYPLLLDVYAGCSKADTVFRLNAT 539
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
588 DRQPKVEYRDIEIDDDNLPMLKLPATFTDTTHYPLLLVVDGTPGSGVAEKFEVSMET 647
Qy 540 YLASTENIIVASFGRGSGYQGDIMHAINRRLGTTFEVEDOIEAARQFSKMGFVNDKRIA 599
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
648 VMVSHGAVVVKCDGRSGFGQTKLLEHVRRLGLLEEKQMEAVRMLKEQYIDRTKVA 707
Qy 600 IWGWSYGGYVTSMLV---GSGSG-VFKCGIAPVPSRWEYDVSVYTRYMGLPTPEDNLD 655
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
708 VFGDYGYLSTYILPAKGENQGTFTCGSALSPTDFKLYASAFSEYRGL---HGLD 763
Qy 656 H--YRNSTVMGRARENFKQVEYLLIHGTADDNVHFPQQAISKALVDVGVDFQAMWYTD 713
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
764 NRAYEMTKVAHRVSALREQQFLIHPHTADEKIHFOHTAELITQLIRGKANYSLQIYPDES 823
Qy 714 HGIASSTAHQHIYTHMSHFIKQCPSL 739
Db 824 HYFTSSSLKQHLRYSIINFFVECPRI 849
```

RESULT 7

```
A41793
dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-s - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41793
R:Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wenthold, R.J.; Shimasaki, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992
A>Title: Differential expression of two distinct forms of mRNA encoding members of a
A:Reference number: A41793; MUID:92108018; PMID:1729689
A:Accession: A41793
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-803 <WAD>
A:Cross-references: UNIPROT:P42659; UNIPARC:UPI0000002A83C; GB:M76429; NID:g408719; PID:
C:Note: sequence extracted from NCBI backbone (NCBIP:75138)
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein
F:257,342/binding site: carbohydrate (Asn) #status predicted
```

```
Query Match 27.5%; Score 1106; DB 2; Length 803;
Best Local Similarity 32.4%; Pred. No. 5.1e-67;
Matches 242; Conservative 158; Mismatches 300; Indels 48; Gaps 22;
```

```
Qy 14 RKTVTLTDYLNKTVRLKLYSLRWISDHYLYK-QENNILVFNARYGNSSVPLENSTPDEF 72
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
66 KKKVTVEDLFSDFDKIHPEAKWISDTEFYREQGTVRLNWNVTNSTVLIEGKKIESL 125
Qy 73 GHSINDYSISDGGFILLNYYVQWRHSYTSASDIYDL--NKRQLITEERIPN-NQOW 129
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
126 -RAIR-YEISPDREYALFSNVNVEPIYQHSYTGYYVLSKIPHGDPQSLDPPPEVSNAKLOYA 183
Qy 130 TWSVPVGHKLAVVWNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWWYEEVFSAYSALW 189
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
184 GWGPKGQQLIFIPENNIYYCAHVCKQAIRVVVSTGKEGVYINGLSDWLYEEBILKTHIAHW 243
Qy 190 WSPNGTFLAYAFNDTEVPLEIYSFYSDESIQLYPKTVVPYPKAGAVNPTVFFVNTDS 249
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
244 WSPDGTRLAVATINDSRVPIWELPTYTSG--VYPTAKPYHPYKAGCENPISLHVI---G 298
```

QY 250 LSSVTNATSIQTAPASMLIGDHYLCDVTWATOERISLQWRRIONYSVMDICDYDESSG 309
DB 299 LNPGRH--DLEMTTPDDPRMRREYIITWVKWATSTKVVNVLGRAQNVSLTLC--DATTG 354
QY 310 RWNCLVARQHIEMTSTGWGRFRSPSPHPTLDCNSFYKIIISNEEG----YRHICYF---- 361
DB 355 ----VCTKKG-EDESAWLHR-QNEPVFSKGRKFFVYRAIPQGGQCKFYHITVSSSQP 408
QY 362 QIDKDCCTPITKGTWEVIGIALTSD-----YLYISNEYKMGPGGRNLYKIQLSDYTKVT 417
DB 409 NSSNDNIQISITGDMVDVTKI--LSYDEKRSQIYFLSTE--DLPRRRQLYSASTVGSFNRQ 464
QY 418 CLSCELNPERCQYYSVSFSKEAKYQLRCSGCLPLTYTLHSSVNDKGLRVLDNSALDKM 477
DB 465 CLSCDL-VNCTYFSAFSGADFFLLKCEGPGVPTVSVHNTTDDKCKMPDLETNEHVQKA 523
QY 478 LQNVOMPSKKLDFIILNETKFWQMLPPHFKSKKYPILLDVYAGPCQKADTVFRLNM 537
DB 524 ISDRQMPKYBYRKIETDDYNIPLQILKPATFTDTAHYPLLLVVDGTPGQSVAEKFAVTM 583
QY 538 ATYLASTENIIVASFDGRSGYQGDKIMHAINRRLGTFVEVDQIEAARFQSKMGFVNDKR 597
DB 584 ETVMVSSHGAUVVKCDGRSGFGQTRLLHEVRRRLGSLBEKQDMEAVRVMLEKPYIDKTR 643
QY 598 IAIWGSYGGVYTSMWL---GSGSG-VFKCGTAVAPVSRWEYDVSVYTERYMGCLPTPEDN 653
DB 644 VAVFGKDYGGYSTYLLPAKGGQAPVFCGSALSPTITDKLYASAFSEYIIGL-----HG 699
QY 654 LDH--YRNTSVMSRAENFKQVEYLLHTGTADONVHFQOQAQISKALVDVGVDFQAMWYTD 711
DB 700 LONRAVEMAKVAHRVSALSQQQFLVIHATADEKIHQHTAELITQLIKKANYSLQIYD 759
QY 712 EDHGIASSTAHOHYIYTHMSHFTKQCFSL 739
DB 760 ESHYFSSAQAQHLHRSILGFFVECPRI 787

RESULT 8
T41703
dipeptidyl aminopeptidase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T41703
R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z22011
A:Accession: T41703
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-793 <MUR>
A:Cross-references: UNIPROT:O14073; UNIPARC:UPI000013AB95; EMBL:AL031180; PIDN:CAA20138.
A:Experimental source: strain 972h; cosmid c2E11 -chimeric
C:Genetics:
A:Gene: SPAC2E11.08
A:Map position: 1
C:Superfamily: dipeptidyl-peptidase IV

Query Match 23.6%; Score 950.5; DB 2; Length 793;
Best Local Similarity 30.5%; Pred. No. 1.8e-56;
Matches 237; Conservative 141; Mismatches 289; Indels 109; Gaps 25;

QY 14 RKTYYTLDYLNTRYLKLKSLRWISDH----EVLKYQENNILVFNABYGNSSVF-LENST 68
DB 52 RDSLSLDDIVLQKYKPSYQVNWIDSGQLKDTPLVKYGLNINIQDPYLNKTLFYSVSLV 111
QY 69 EDEFSGHSINDYSTSPGQFLLTYVYVVKWRHSHYSATSYDIYDLNKLQLITEIRPN--- 124
DB 112 INGIQLDYDYSISLSPFAKTVLVSNKSRWRHSSFAQYLYDN-----TETKDVNMLQ 164
QY 125 -NTQW----VTVSPVGHKGLAYVWNNDIYV-KIEPNLPYSRITWTGKEDIYNGITDWVYE 178
DB 165 DNEHWTISLAEWSPGTGQLSFVYNNDLYVRKNDGNV--QRLTYDGTVD-VFNGLTDWIYE 221

QY 179 BEVFSAYSALWWSPNGTFLVAQPNDETVPLIEYSFYSGE-----SLQYPKTVRVYPYK 232
DB 222 BEVLSSPTIWNWSPDSDKIAFLKLNSEIPIYHYPLTYAELDPSLPEDFYNDKMAIKYK 281
QY 233 AGAVNPTVKFFVYVNTDLSLSSVTNATSIQTAPASMLIGDHYLCDVTWATQERISIQWLR 292
DB 282 PCGNPNPVSFLV--ADLNSNASSNPSLWNEP---LAEPVQVQVNLVWNTSSVLVQFTNR 335
QY 293 IONYSVMDICDYDESSGRWNCCLVARQHIEMTSTGWGRFRSPSPHPTLD-----GNS 344
DB 336 NSTCITABLLDTELKS-----IHTVKYTECLEBEGY-EVQOQAKMFFPLNLSLVNWNWSDG 388
QY 345 FYKIIISNEEGYRHICYFQIDKDCCTPITKGTWEVIG-----IEALTSD 387
DB 389 YFDIILALDD-YNHLAFIPFGSSPIYLTSGAWDVTDPGHIDGDFGNVYFLATLKDSTER 447
QY 388 YLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQLRCS 447
DB 448 HLYYVS-----LDTLEIYGITD-----NGEDEGYYSTSFSPFGDFVYLYNH 488
QY 448 GPGLPLYTLHSSVNDKGLRVLEDNSALOKMLQNVOMPSKKLDFIILNETKFWQMLLPH 507
DB 489 GPDVPMQELRSTKOKDYCLSLSTNSRLKQOLSSITLPSVEYQKLTFTDFTFNMERRPRN 548
QY 508 FDKSKKYPLLLDVYAGPCQKADTVFRLNWTATYLASTEN--IIVASFQDGRSGYQGDKIM 565
DB 549 FQVKKYYPVLFPAYGPGSQQVAKLFRVDVFNQYLAHSHDPFPIVTLDRGTGFGNAPR 608
QY 566 HAINRRLGTFEVEDQIEAARFQSKMGFVNDKRKRIAIWGSYGGVYTSMWLGSQGVFKCGI 625
DB 609 YSVSRHLGEWSYDQOQAGKFWADLPFVDENHVGWGSYGGYLLTKLET-QDVFSGM 667
QY 636 AVAPVSRWEYDVSVYTERYMGCLPTPEDNLDHYRNTSVMSRAENFKQVE-YLLIHGTADDN 684
DB 668 AVAPVTDWRLYDSVYTERYMDL--PQYNKEGYKNSQIHD-YBKQQLKRFVFAHGFGDDN 724
QY 685 VHFQOQAQISKAL-----VDVGVDFQAMWYTDDEHGIASSTAHOHYIYTHMSHFI 733
DB 725 VHFQSHMLMDGLNLANCYNVDMV-----FPDSAHSISYHNASLSIYHRUSEWI 774

RESULT 9
A30107
dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protease YHR028C
C:Species: Saccharomyces cerevisiae
C:Date: 07-Jun-1990 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C:Accession: S46780; A30107
R:Du, Z.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 8082.
A:Reference number: S46773
A:Accession: S46780
A:Molecule type: DNA
A:Residues: 1-818 <DUZ>
A:Cross-references: UNIPROT:P18962; UNIPARC:UPI0000031A5F; EMBL:U10399; NID:g500689; PTI
R:Roberts, C.J.; Pohlrig, G.; Rothman, J.H.; Stevens, T.H.
J. Cell Biol. 108, 1363-1373, 1989
A:Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an i.
A:Reference number: A30107; MUID:89174971; PMID:2647766
A:Accession: A30107
A:Molecule type: DNA
A:Residues: 1-82, 'H', 84-124, 'N', 126-181, 'LRRLET', 189-199, 'N', 201-365, 'DFKRGKERKE', 376-571
A:Cross-references: UNIPARC:UPI0000172A3C; EMBL:X15484
A>Note: the authors translated the codon ACC for residue 572 as Asn
C:Genetics:
A:Gene: SGD:DAP2; STE13; MIPS:YHR028C
A:Cross-references: SGD:S0001070; MIPS:YHR028C
A:Map position: 8R
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidyl-peptide hydrolase; glycoprotein; transmembrane protein; yeast vacu
F:30-45/Domain: transmembrane #status predicted <TMM>

RESULT 10
S66261
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum
C.Species: Flavobacterium meningosepticum
C.Date: 28-Oct-1995 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
R.Accession: S66261
R.Kabashima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
Arch. Biochem. Biophys. 320, 123-128, 1995
A.Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV
A.Reference number: S66261; MUID:95314307; PMID:7793970

RESULT 11
T19514

hypothetical protein C27C12.7 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #taxid 6233
 C:Accession: T19514
 R:Thomas, K.
 submitted to the EMBL Data Library, March 1996

[illegible]

RESULT 13
T25174
hypothetical protein T23F1.7b - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T25174
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z19990
A/Accession: T25174
A/Status: preliminary; translated from GB/EMBL/DDBBJ
A/Molecule type: DNA
A/Residues: 1-799 <MIL>
A/Cross-references: UNIPROT:O18119; UNIPARC:UPI0000061190; EMBL:Z81129; PIDN:CAB03412.1
A/Experimental source: clone T23F1
C/Genetics:
A/Gene: CRSP:T23F1.7b
A/Map position: 5
A/Introns: 13/3; 52/1; 111/3; 197/3; 366/3; 524/1; 557/1; 648/2; 703/1; 743/1
C/Superfamily: dipeptidyl-peptidase IV

[illegible]

[illegible]

RESULT 15
A87516
dipeptidyl peptidase IV [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: A87516
C:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolodziej, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A>Title: Complete Genome Sequence of *Caulobacter crescentus*.
A|Reference number: A87249; MUID:21173698; PMID:11259647
A|Accession: A87516
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-738 <STO>
A|Cross-references: UNIPROT:Q9ASE0; UNIPARC:UPI00000C7616; GB:AE005673; NID:G13423647; PDB:1G83
C|Genetics:
A|Gene: CC2154

Query Match	15.64;	Score 628;	DB 2;	Length 738;
Best Local Similarity	26.84;	Pred. No. 1.2e-34;		
Matches	189;	Conservative 115;	Mismatches 269;	Indels 132; Gaps 23
Qy	76	INDYSIPDQFILLLEYN--YVQWRHSYTSYDIVDLNKRQLITEERIPNNTQWTVS	132	
Db	111	IVEYSWRQGRFILLVPLDGDLYL-----DAVADGKILTRLE--TPGDEVDKVS	157	
Qy	133	PVGHKLAYVWNNDIYVRIEPLPSYRITWTGKBDIIYNGITDWMYEEVFSAYSALMWSP	192	
Db	158	PKGQVSVYVRDQNLKYPVAGGAETALTDDCKDALSFGVAEFVQEBEL-DRFTGYWVNSP	215	
Qy	193	NGTFLAYAQFNDTEVPLIEVSFYDSLSQYFKTVRVVPYKPGAGANVPTKFFVWVTDLSLS	252	
Db	216	DESRIVTVRDESGVDIVP---RADIGFGGATVNVQRYPRAGRPNNAVVDLP-----	263	
Qy	253	VYNATSIQITAPASMLIGDHYLDCDVTW-ATQERISLOWLRIONYSVMDICDYEDESSGRW	311	
Db	264	VRDLASGKVVALDGLAKNDIYVARVMSADGKTVVQRLSRDQ--KTLDLLAFDAATGAG	321	
Qy	312	NCLVARQHIEMSTTGWGRFPSPSEPHFTLDGNSFYKII-----SNBEGYRHICYFQID	364	
Db	322	KTILT-----DTPDHFIEVSNDPRPLTDGTFLWGSEKDGNOHLYRYAAD	365	
Qy	365	KQDCTFITKGTWEVIGIEALTSVDLYIYSNBYKMGPGGGRNLYKIQLSDTYKTCISCELN	424	
Db	366	GKLIATIKTGDWPVIGLEG-----VDEARKVAIFESASID	399	
Qy	425	-PERCQVYSYFSKEAKYQQLRCSG-----PGLPLYTLHSSVNDK	463	
Db	400	TPIBRRLYEVSYAKPGPKALTSAGGWAAKVAONGGAPAGTYSDPKTPSQOTALYSADGK	459	
Qy	464	GLRVLEDNSALD-----KMLQNVQMS-----KCLDFFILNETKFWYQMLPLPHFPDKSKY	514	
Db	460	RVRWIEENKLAEGHPYYPYAAANLQPEFGSLKAAD---GET-LHYEILKPIGDFPAKKY	514	
Qy	515	PLLLDVYAGPCSQKADTVFRLNW-----ATYLASTENIIVASFQDGGSGYQDKMKHAIN	569	
Db	515	PAIVSVYGGPHQARV---MKQWHSYSPSRITYLEA--GYVIFKLDNRNGSNRSKAFPMRALD	568	
Qy	570	RLRGTFEVEDQIEARQFSKMGFVDNKRKRIATWGSYSGVVTSMVLGSSGVFKCGIYAVAP	629	
Db	569	RKLGTFEVEDQLLGAFLASQPYVDADKLGVGWGSYSGFMAULMLTAENTPPKAGAGAP	628	
Qy	630	VSRWEYDYSVTVRYMGLTPEDNLIDHYNRNSTWMSRAENFKQVEYLLIHGTADDNVHFOQ	689	
Db	629	PTEHSLVDYATERYMG--KPDENKAGYAYSDINNRIKDLAPGSLLLHGMADDNVIFEN	686	
Qy	690	SAQISKALVDVGVDVQAMWYTDDEHGIASSTAH--QHIYVTWMSHF	732	
Db	687	STRMLAALORKAILFENAMYGGERHSAPGSKTKGLSVLTKHLDDFF	731	

Search completed: February 15, 2006, 13:15:27
Job time : 48 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 13:26:39 ; Search time 18 Seconds
(without alignments)
584.259 Million cell updates/sec

Title: US-10-659-055-3

Perfect score: 4020

Sequence: 1 ADPGSHHHHRSKTYTLT.....AQHIYTHMSHFHKQCSLP 740

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 107799 seqs, 14211699 residues

Total number of hits satisfying chosen parameters: 107799

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB pep.*
- 5: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB pep.*
- 7: /cgn2_6/ptodata/2/pubaa/US11_NEW_PUB pep.*
- 8: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3939	98.0	766	6	US-10-522-789-2
2	3886	96.7	762	7	US-11-116-939-13
3	2168	53.9	760	7	US-11-186-284-55
4	550	13.7	102	7	US-11-176-951-7
5	528.5	13.1	882	7	US-11-151-601-20
6	410.5	10.2	102	7	US-11-176-951-8
7	291	7.2	54	7	US-11-179-977-1
8	272	6.8	657	7	US-11-176-951-5
9	167	4.2	115	7	US-11-176-951-12
10	167	4.2	115	7	US-11-176-951-12
11	145	3.6	27	7	US-11-176-951-6
12	145	3.6	27	7	US-11-176-951-13
13	145	3.6	109	7	US-11-176-951-10
14	130	3.2	99	7	US-11-176-951-11
15	124	3.1	668	6	US-10-454-437-118
16	116.5	2.9	1593	6	US-10-453-372-1092
17	114	2.8	960	7	US-11-169-041-177
18	114	2.8	2516	6	US-10-647-956A-2
19	112.5	2.8	1436	6	US-10-453-372-1094
20	110	2.7	19	7	US-11-116-939-27
21	110	2.7	72	7	US-11-151-601-22
22	110	2.7	453	7	US-11-024-959-355
23	109.5	2.7	877	7	US-11-077-550-157
24	107.5	2.7	3194	7	US-11-052-554A-90
25	107	2.7	866	7	US-11-077-550-32

26	107	2.7	871	7	US-11-077-550-155	Sequence 155, App
27	107	2.7	873	7	US-11-077-550-163	Sequence 163, App
28	107	2.7	873	7	US-11-077-550-165	Sequence 165, App
29	107	2.7	876	7	US-11-077-550-66	Sequence 66, Appl
30	107	2.7	876	7	US-11-077-550-62	Sequence 62, Appl
31	107	2.7	879	7	US-11-077-550-30	Sequence 30, Appl
32	107	2.7	908	7	US-11-077-550-64	Sequence 64, Appl
33	107	2.7	914	7	US-11-077-550-60	Sequence 60, Appl
34	107	2.7	949	7	US-11-077-550-68	Sequence 68, Appl
35	107	2.7	1130	7	US-11-077-550-44	Sequence 44, Appl
36	107	2.7	1132	7	US-11-077-550-46	Sequence 46, Appl
37	107	2.7	4495	6	US-10-453-372-1002	Sequence 1002, Ap
38	106	2.6	878	7	US-11-077-550-12	Sequence 12, Appl
39	106	2.6	907	7	US-11-077-550-16	Sequence 16, Appl
40	106	2.6	953	7	US-11-077-550-14	Sequence 14, Appl
41	106	2.6	1013	7	US-11-077-550-18	Sequence 18, Appl
42	105.5	2.6	305	7	US-11-098-686-10980	Sequence 10980, A
43	105.5	2.6	658	6	US-10-873-528-17	Sequence 17, Appl
44	105.5	2.6	677	6	US-10-873-528-155	Sequence 155, App
45	105.5	2.6	871	7	US-11-077-550-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-522-789-2

; Sequence 2, Application US/10522789

; Publication No. US20050260732A1

; GENERAL INFORMATION:

; APPLICANT: TANABE SEIYAKU CO., LTD.

; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV

; FILE REFERENCE: 03-039-PCT

; CURRENT APPLICATION NUMBER: US/10/522,789

; CURRENT FILING DATE: 2005-01-28

; PRIOR APPLICATION NUMBER: US 60/398,761

; PRIOR FILING DATE: 2002-07-29

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 766

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-522-789-2

Query Match 98.0%; Score 3939; DB 6; Length 766;

Best Local Similarity 100.0%; Pred. No. 2.2e-298;

Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	13	SRKTYTLTDYLNKNTYRLKLSLRWISDHEVLYKQENILVFNABYGNSSVPLENSTPDEF	72
Db	39	SRKTYTLTDYLNKNTYRLKLSLRWISDHEVLYKQENILVFNABYGNSSVPLENSTPDEF	98
Qy	73	GHSINDYSISPDGQFILLENNYKQWRHSYTSYDIYDLNKRQLITEIRIPNNTQWTVWS	132
Db	99	GHSINDYSISPDGQFILLENNYKQWRHSYTSYDIYDLNKRQLITEIRIPNNTQWTVWS	158
Qy	133	PVGHKLAVVWNNDIYVKIEPNLPSYRLTWGKEDIYNGITDWWYEEVPSAYSALWSP	192
Db	159	PVGHKLAVVWNNDIYVKIEPNLPSYRLTWGKEDIYNGITDWWYEEVPSAYSALWSP	218
Qy	193	NGTFPLAQAQNDIEVPLIEYSFYSDESLOQPKTVRPVPKAGAVNPTVKFVWNTDSLSS	252
Db	219	NGTFPLAQAQNDIEVPLIEYSFYSDESLOQPKTVRPVPKAGAVNPTVKFVWNTDSLSS	278
Qy	253	VTNATSIQTAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN	312
Db	279	VTNATSIQTAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN	338
Qy	313	CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNBEGRHICVFQIDKKDCTFIT	372
Db	339	CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNBEGRHICVFQIDKKDCTFIT	398

Qy	373	KGTEVIGIEALTSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYVS	433
Db	399	KGTEVIGIEALTSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYVS	458
Qy	433	VSPKEAKYYQLRCGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDFII	492
Db	459	VSPKEAKYYQLRCGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDFII	518
Qy	493	LNETHFWYQMLTLPHPFDKSKKYPILLDYYAGPCOKADTVFRLNWTATYLASTENIIVASF	552
Db	519	LNETHFWYQMLTLPHPFDKSKKYPILLDYYAGPCOKADTVFRLNWTATYLASTENIIVASF	578
Qy	553	DGRSGYGQDKIMHAINRRLGTFEVEDQIEAARQSKMGFVDNKRKIAIWGMSYGGVVTSM	612
Db	579	DGRSGYGQDKIMHAINRRLGTFEVEDQIEAARQSKMGFVDNKRKIAIWGMSYGGVVTSM	638
Qy	613	VLGSGGVFKCGIAVAPVSRWEYDVSVTERYMGILPTPEDNLHDYRNSTVMSRAENFKQV	672
Db	639	VLGSGGVFKCGIAVAPVSRWEYDVSVTERYMGILPTPEDNLHDYRNSTVMSRAENFKQV	698
Qy	673	EYLLIHGTADDNVHFQOSAQSICALVDGVDPQAMWYTDEDHGIASSTAHQHIYTHMSHF	732
Db	699	EYLLIHGTADDNVHFQOSAQSICALVDGVDPQAMWYTDEDHGIASSTAHQHIYTHMSHF	758
Qy	733	IKQCFSLP 740	
Db	759	IKQCFSLP 766	
RESULT 2			
US-11-116-939-13			
; Sequence 13, Application US/11116939			
; Publication No. US20050265995A1			
; GENERAL INFORMATION:			
; APPLICANT: Stephen Tomlinson			
; APPLICANT: Richard J. Quigg			
; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS			
; FILE REFERENCE: 19113.0115U2			
; CURRENT APPLICATION NUMBER: US/11/116,939			
; CURRENT FILING DATE: 2005-04-28			
; PRIOR APPLICATION NUMBER: 60/565,907			
; PRIOR FILING DATE: 2004-04-28			
; NUMBER OF SEQ ID NOS: 27			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 13			
; LENGTH: 762			
; TYPE: PRT			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic			
; OTHER INFORMATION: construct			
US-11-116-939-13			
Query Match 96.7%; Score 3886; DB 7; Length 762;			
Best Local Similarity 99.0%; Pred. No. 2.8e-294;			
Matches 721; Conservative 0; Mismatches 3; Indels 4; Gaps 1			
Qy	13	SRKTYTLTDYLNKTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVPLENSTDFEF	72
Db	39	SRKTYTLTDYLNK-----YLNKCLRWISDHELYLKQENNILVFNAEYGNSSVPLENSTDFEF	94
Qy	73	GHSINDYSISDPGQFILLLEYNVYKQWRHSYTSASYDIYDLNKKQLITEERIINNTOVWTS	132
Db	95	GHSINDYSISDPGQFILLLEYNVYKQWRHSYTSASYDIYDLNKKQLITEERIINNTOVWTS	154
Qy	133	PVGHKLAIVWVNDIYVKLEPNLPSYRIIWTGKEDIYNGITDWTVEEYVFSAYSALWWSF	192
Db	155	PVGHKLAIVWVNDIYVKLEPNLPSYRIIWTGKEDIYNGITDWTVEEYVFSAYSALWWSF	214
Qy	193	NGTFLAYAQFNDTEVPLIEYGFYSDESIOYKPTVRVPYKAGVNPVKFVVNTDSLSS	252
Db	215	NGTFLAYAQFNDTEVPLIEYGFYSDESIOYKPTVRVPYKAGVNPVKFVVNTDSLSS	274


```
QY 54 NAEYG-----NSSVFLNSTFDFGHSINDYSISPDQOFILLFYNVYKQWRHSYTSYDIY 109
Db 126 TLDYGMYSREBLLRKRKIGTVGIASYDYH-QSGTFL-----FQAGSGIY 171
QY 110 DLNK-----RLTTEIRIPN-----NTQVTVSPVGHKLAYVNNDIYV 148
Db 172 HVKGGPQGFQTOPLRPNLVETSCNIRMDPKLCPADPMI-----AFHSNDIWI 222
QY 149 KIEPNLPSVITWTG-----KEDIYNGITWVVEEVFSAYSALMWSPNQTF-----196
Db 223 SNIVTREERRLTYVHNELANBEDARSAGVATFVLQEE-FDRYSGYVWCPEKAETTPSGGK 281
QY 197 ---LAAQPNDETEVPLIEYSFSDLESLOYPKTVRPYPKAGAVNPVKEFV---VNTDSLS 251
Db 282 ILRIIYEENDESEVEIHH---VTSPLMETRADSPRYPKTGANTPKVTFPMSEIMIDABG 338
QY 252 SVTNATSIQITAPASMLI-GDHYLCDVTWA-----TOERISLOWL-----290
Db 339 RIIDVIDKELIQPFELFEGVEYIARAGWTPGKYAWSILLDRSQRLQIVLISPFLFP 398
QY 291 -----RRTONY---SVMDICDVDESSGRW-NC-----LVARQHIEMSTTGWGRRFP 333
Db 399 VEDDVMERQLIESVPDSVTPLIYEETTDIWINIHDFHVPFQSGHEEIEBIFASECKT 458
QY 334 SEPHTLDGNSFYKILS---NEBGRYHI-----CYFQIDKKDCTFITKGTWEVIG---380
Db 459 GFRH-----LYKITSILKSKYKRSKSGGLPAPSDFKCPKIBEEIAITSGEWEVLGRHGS 511
QY 381 -TEALTSYLYISNEYKMPGGRNLYKI-----QLSD---YTKVTCLSCELNPERC 428
Db 512 NQOVDEVRLVYFEGT-KDSPLEHLYVYVYNPGEVTLTRDGYSHSCIS-----QHC 565
QY 429 QYYSVSFSKAKYQLRCSGGLPLTYLHSSVNDKGLRVLED-NSALDKM---LQNVQWPS 485
Db 566 DPFISKYSQKNPH---C---VSLYKLSPPDDPTCKTKFETWATILDSAGPLPDVTPPE 618
QY 486 KKLDFILLNETXF-WYQMLLPHP-PDKSKYPLLLDVVAGPCSQ-----KADTVFRLNW 537
Db 619 ---IFSFESTTGTFLYGLMYLKHDLQPGKKYPTVLFIYGGPQVLVNNRFGVKYFRLN- 674
QY 538 ATYLASTENIIVASFGRSGYQGDKIMHAINRRLGTFEVEQIEAARQF-SKMGFVDNK 596
Db 675 --TLASLGIVVVV-IDNRSGCHRLKFECAFYKMGQIEIDQVEGLVLAIRYDFIDL 731
QY 597 RIAIWMWSYGGVYTSVNLGSGGVFKGIAVAPVSRWEYVDSYTYRYNGLPTPTPDNLDH 656
Db 732 RVGIHWSYGGVLSLMLAQMSDIFRAVALAGAPVTLWIFDYGTYERYMG--HPDQNEQG 789
QY 657 YRNSTVMSAENF--KQVEYLLIHGTADDNVHVFQOQAISKALVDVGVDFQAMWYTDREH 714
Db 790 YVLGSVMAQAEKPPSEFNRLHLHGLDENVHFHFAHTSILLSLFVLRAGKPYDILQIYPOERH 849
QY 715 GIASSTAHOHIYTHMSHFIFIK 735
Db 850 SIRVPESGEHYELHLLHYLQE 870

RESULT 6
US-11-176-951-8
; Sequence 8, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952

Query Match 7.2%; Score 291; DB 7; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 687 FQOQAISKALVDVGVDFQAMWYTDREHGHASSTAHOHIYTHMSHFIFIKQCFSLP 740
Db 1 FQOQAISKALVDVGVDFQAMWYTDREHGHASSTAHOHIYTHMSHFIFIKQCFSLP 54

RESULT 8
US-11-179-977-1
; Sequence 1, Application US/11179977
; Publication No. US20050249789A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
; FILE REFERENCE: GC511-PCT
; CURRENT APPLICATION NUMBER: US/11/179,977
; CURRENT FILING DATE: 2005-07-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Bacillus
US-11-179-977-1
```

```
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 8
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-8

Query Match 10.2%; Score 410.5; DB 7; Length 102;
Best Local Similarity 72.3%; Pred. No. 2.7e-25;
Matches 73; Conservative 17; Mismatches 10; Indels 1; Gaps 1;

QY 639 VYTRYMGLPPTEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHVFQOQAISKALV 698
Db 1 VYTRYMGLPPTEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHVFQOQAISKALV 60

QY 699 DVGVDFQAMWYTDREHGHASSTAHOHIYTHMSHFIFIKQCFSL 739
Db 61 NAQVDFQAMWYSDQNHGL-SGLSTNHLTYTHMTHFLKQCFSL 100
```

```
RESULT 7
US-11-176-951-5
; Sequence 5, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 5
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-5
```

```
Query Match 7.2%; Score 291; DB 7; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 687 FQOQAISKALVDVGVDFQAMWYTDREHGHASSTAHOHIYTHMSHFIFIKQCFSLP 740
Db 1 FQOQAISKALVDVGVDFQAMWYTDREHGHASSTAHOHIYTHMSHFIFIKQCFSLP 54
```

```
RESULT 8
US-11-179-977-1
; Sequence 1, Application US/11179977
; Publication No. US20050249789A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
; FILE REFERENCE: GC511-PCT
; CURRENT APPLICATION NUMBER: US/11/179,977
; CURRENT FILING DATE: 2005-07-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Bacillus
US-11-179-977-1
```


; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 6
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-6

Query Match 3.6%; Score 145; DB 7; Length 27;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 PLIEYSFYSDLSQYPKTVRVYPKAG 234
Db 1 PLIEYSFYSDLSQYPKTVRVYPKAG 27

RESULT 12
US-11-176-951-13
; Sequence 13, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; PRIOR FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 13
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-13

Query Match 3.6%; Score 145; DB 7; Length 27;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 PLIEYSFYSDLSQYPKTVRVYPKAG 234
Db 1 PLIEYSFYSDLSQYPKTVRVYPKAG 27

RESULT 13
US-11-176-951-10
; Sequence 10, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 10
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens

US-11-176-951-10

Query Match 3.6%; Score 145; DB 7; Length 109;
Best Local Similarity 33.7%; Pred. No. 0.00013;
Matches 33; Conservative 17; Mismatches 44; Indels 4; Gaps 2;

QY 640 YTERYMGLPPTDNLHDYRNSTVMSRAENF--KQVEYLLIHGTADDNVHVFQSOAQISKAL 697
Db 2 YTERYMG--HPDNEQGYLGSVAMQAEKFPSEPNRLLLHGFLENVHVFHTNPLVSQL 59
QY 698 VDVGVDFQAMMYTDEDHGIASTAHQHIYTHMSHFQK 735
Db 60 VRAGKPYQLQIYPQSRHSIRVPESGHEVELHLLHFLQE 97

RESULT 14

US-11-176-951-11
; Sequence 11, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-11

Query Match 3.2%; Score 130; DB 7; Length 99;
Best Local Similarity 30.6%; Pred. No. 0.0017;
Matches 30; Conservative 18; Mismatches 46; Indels 4; Gaps 2;

QY 640 YTERYMGLPPTDNLHDYRNSTVMSRAENF--KQVEYLLIHGTADDNVHVFQSOAQISKAL 697
Db 2 YTERYMDV--PENNOGHYEAGSVLHVKEKLPNEPNRLLLHGFLENVHVFHTNPLVSQL 59
QY 698 VDVGVDFQAMMYTDEDHGIASTAHQHIYTHMSHFQK 735
Db 60 IRAGKPYQLQIYPNERHSIRCPESGHEVEVTLHFLQE 97

RESULT 15

US-10-454-437-118
; Sequence 118, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberkauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 13:25:59 ; Search time 168 Seconds
(without alignments)
1840.439 Million cell updates/sec

Title: US-10-659-055-3

Perfect score: 4020

Sequence: 1 ADPGSHHHHSRKYTLT.....AQHIYTHMSHFQKQSLP 740

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_Main:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3939	98.0	766	3	US-09-265-606-3
2	3939	98.0	766	3	US-09-993-959-1
3	3939	98.0	766	4	US-10-295-027-590
4	3939	98.0	766	5	US-10-952-459-18
5	3939	98.0	766	5	US-10-631-457-565
6	3933	97.8	766	4	US-10-002-593-6
7	3933	97.8	766	4	US-10-165-603-7
8	3933	97.8	766	4	US-10-423-714-6
9	3933	97.8	766	4	US-10-295-027-922
10	3933	97.8	766	4	US-10-794-899-41
11	3933	97.8	766	6	US-11-041-674-6
12	3929	97.7	766	5	US-10-476-264-147
13	3929	97.7	766	5	US-10-770-712-1
14	3406.5	84.7	767	5	US-10-770-712-2
15	3395.5	84.5	767	4	US-10-165-603-4
16	3395.5	84.5	767	4	US-10-794-899-38
17	3390	84.3	760	5	US-10-631-467-1390
18	3374	83.9	760	5	US-10-770-712-3
19	2175	54.1	760	5	US-10-723-860-4171
20	2168	53.9	760	3	US-09-265-606-2
21	2168	53.9	760	4	US-10-177-293-136
22	2168	53.9	760	4	US-10-301-822-55
23	2168	53.9	760	5	US-10-884-070A-13
24	1289.5	32.1	504	4	US-10-072-012-863
25	1223	30.4	746	5	US-10-476-264-69
26	1223	30.4	746	5	US-10-476-264-70
27	1223	30.4	746	5	US-10-476-264-72

28	1223	30.4	789	5	US-10-476-264-71	Sequence 71, Appl
29	1223	30.4	796	3	US-09-976-674-5	Sequence 5, Appl
30	1223	30.4	796	3	US-09-870-133-2	Sequence 2, Appl
31	1223	30.4	796	4	US-10-160-501-5	Sequence 5, Appl
32	1223	30.4	796	5	US-10-982-512-5	Sequence 5, Appl
33	1223	30.4	796	5	US-10-433-757-2	Sequence 2, Appl
34	1223	30.4	796	5	US-10-476-264-68	Sequence 68, Appl
35	1223	30.4	796	5	US-10-476-264-106	Sequence 110, App
36	1223	30.4	796	5	US-10-476-264-110	Sequence 114, App
37	1223	30.4	796	5	US-10-476-264-142	Sequence 142, App
38	1223	30.4	796	5	US-10-476-264-146	Sequence 146, App
39	1223	30.4	797	5	US-10-476-264-36	Sequence 36, Appl
40	1221	30.4	228	4	US-10-476-264-114	Sequence 114, App
41	1217	30.3	798	4	US-10-210-130-120	Sequence 120, App
42	1198	29.8	743	4	US-10-363-937-4	Sequence 4, Appl
43	1198	29.8	743	6	US-11-167-040-4	Sequence 4, Appl
44	1196	29.8	706	3	US-09-976-674-41	Sequence 41, Appl
45	1196	29.8	706	5	US-10-982-512-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-09-265-606-3
; Sequence 3, Application US/09265606
; Patent No. US20020034789A1
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020034789A1man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-265-606-3

Query Match 98.0%; Score 3939; DB 3; Length 766;

Best Local Similarity 100.0%; Pred. No. 3,7e-306; Indels 0; Gaps 0;
Matches 726; Conservative 0; Mismatches 0;

Qy 13 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENILVFNABYGNISVFLENSTFDEF 72

```

Db      39 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDEF 98
      73 GHSINDYSIPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 132
      99 GHSINDYSIPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 158
      133 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWWYEEVEFSAYSALWWS 192
      159 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWWYEEVEFSAYSALWWS 218
      193 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLS 252
      219 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLS 278
      253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN 312
      279 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN 338
      313 CLVARQHIEMSTGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 372
      339 CLVARQHIEMSTGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 398
      373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCSELNPERCQYYS 432
      399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCSELNPERCQYYS 458
      433 VSFSKEAKYQYLRCSGPGPLVTLHSSVNDKGLRVLEDSALDKMLQNVQMPSKKLDFTI 492
      459 VSFSKEAKYQYLRCSGPGPLVTLHSSVNDKGLRVLEDSALDKMLQNVQMPSKKLDFTI 518
      493 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTYLASTENIIVASF 552
      519 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTYLASTENIIVASF 578
      553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVGNKRIAIWGNYSYGGYVTSM 612
      579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVGNKRIAIWGNYSYGGYVTSM 638
      613 VLGSAGVFKCGIAPVAPSRWEYDSVYTERYMGVLPEDNLDHYNSTVMSRAENFKQV 672
      639 VLGSAGVFKCGIAPVAPSRWEYDSVYTERYMGVLPEDNLDHYNSTVMSRAENFKQV 698
      673 EYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTTDEHGIASSTAHOIYTHMSHF 732
      699 EYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTTDEHGIASSTAHOIYTHMSHF 758
      733 IKQCFSLP 740
      759 IKQCFSLP 766

```

RESULT 2

```

US-09-993-959-1
; Sequence 1, Application US/0993959
; Publication No. US20030165489A1
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/993,959
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 766
; TYPE: FRT
; ORGANISM: Homo sapiens
US-09-993-959-1

```

Query Match 98.0%; Score 3939; DB 3; Length 766;

```

Best Local Similarity 100.0%; Pred. No. 3.7e-306;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
      13 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDEF 72
      39 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDEF 98
      73 GHSINDYSIPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 132
      99 GHSINDYSIPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 158
      133 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWWYEEVEFSAYSALWWS 192
      159 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWWYEEVEFSAYSALWWS 218
      193 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLS 252
      219 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLS 278
      253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN 312
      279 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN 338
      313 CLVARQHIEMSTGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 372
      339 CLVARQHIEMSTGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 398
      373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCSELNPERCQYYS 432
      399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCSELNPERCQYYS 458
      433 VSFSKEAKYQYLRCSGPGPLVTLHSSVNDKGLRVLEDSALDKMLQNVQMPSKKLDFTI 492
      459 VSFSKEAKYQYLRCSGPGPLVTLHSSVNDKGLRVLEDSALDKMLQNVQMPSKKLDFTI 518
      493 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTYLASTENIIVASF 552
      519 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTYLASTENIIVASF 578
      553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVGNKRIAIWGNYSYGGYVTSM 612
      579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVGNKRIAIWGNYSYGGYVTSM 638
      613 VLGSAGVFKCGIAPVAPSRWEYDSVYTERYMGVLPEDNLDHYNSTVMSRAENFKQV 672
      639 VLGSAGVFKCGIAPVAPSRWEYDSVYTERYMGVLPEDNLDHYNSTVMSRAENFKQV 698
      673 EYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTTDEHGIASSTAHOIYTHMSHF 732
      699 EYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTTDEHGIASSTAHOIYTHMSHF 758
      733 IKQCFSLP 740
      759 IKQCFSLP 766

```

RESULT 3

```

US-10-295-027-590
; Sequence 590, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; Methods of Screening for Modulators of Cancer

```

FILE REFERENCE: 018501-0125000S
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 590
LENGTH: 766
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-590

Query Match 98.0%; Score 3939; DB 4; Length 766;
Best Local Similarity 100.0%; Pred. No. 3.7e-306;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 SRKTYTTLTYLKNTRYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTPDEF 72
Db 39 SRKTYTTLTYLKNTRYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTPDEF 98
Qy 73 GHSINDYSISDGGQFILLENNYVQWRHSYTSYDIYDLNKRQLITEERIINNQTQWTWS 132
Db 99 GHSINDYSISDGGQFILLENNYVQWRHSYTSYDIYDLNKRQLITEERIINNQTQWTWS 158
Qy 133 PVGHKLAVVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSAWMSP 192
Db 159 PVGHKLAVVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSAWMSP 218
Qy 193 NGTFLAYAQFNDTEVPLIEYSFYDESLOYPKTVRVPYPKAGAVNPTVKPFVNTDSLSS 252
Db 219 NGTFLAYAQFNDTEVPLIEYSFYDESLOYPKTVRVPYPKAGAVNPTVKPFVNTDSLSS 278
Qy 253 VTNATSIQITAPASMLIGDHYLCDVWATQERISLOWLRRIONYSVMDICDYDESSGRWN 312
Db 279 VTNATSIQITAPASMLIGDHYLCDVWATQERISLOWLRRIONYSVMDICDYDESSGRWN 338
Qy 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEBGYRHCYFQIDKKDCTFIT 372
Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEBGYRHCYFQIDKKDCTFIT 398
Qy 373 KGTWEVIGIEALTSYDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCISCELNPERCQYIS 432
Db 399 KGTWEVIGIEALTSYDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCISCELNPERCQYIS 458
Qy 433 VSPSKEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFII 492
Db 459 VSPSKEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFII 518
Qy 493 LNETKFWQMLPPHFDKSKYKYPILLDVAAGPCSKADTVFRLNWTATYLASTENIIVASF 552
Db 519 LNETKFWQMLPPHFDKSKYKYPILLDVAAGPCSKADTVFRLNWTATYLASTENIIVASF 578
Qy 553 DGRSGYQGDKIMHAINRRLGTFFVEBQIEAARQFSKMGFVDNKRRIALWGSYGYVTSM 612

Db 579 DGRSGYQGDKIMHAINRRLGTFFVEBQIEAARQFSKMGFVDNKRRIALWGSYGYVTSM 638
Qy 613 VLGSQGVFKCGIAVAPVSRWEYDVSYYTERYMGILPTPEDNLDHYRNSTVMSRAENPKQV 672
Db 639 VLGSQGVFKCGIAVAPVSRWEYDVSYYTERYMGILPTPEDNLDHYRNSTVMSRAENPKQV 698
Qy 673 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAWWYTDDEHGFIASSTAHQHIYTHMSHF 732
Db 699 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAWWYTDDEHGFIASSTAHQHIYTHMSHF 758
Qy 733 IKQCFSLP 740
Db 759 IKQCFSLP 766

RESULT 4
US-10-952-459-18
Sequence 18, Application US/10952459
Publication No. US20050074805A1
GENERAL INFORMATION:
APPLICANT: Kochan, Jarema Peter
APPLICANT: Martin, Mitchell Lee
APPLICANT: Robinski, James Andrew
TITLE OF INVENTION: Specific Markers for Diabetes
FILE REFERENCE: 21270US1
CURRENT APPLICATION NUMBER: US/10/952,459
CURRENT FILING DATE: 2004-09-28
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patent in version 3.2
SEQ ID NO 18
LENGTH: 766
TYPE: PRT
ORGANISM: Homo sapiens
US-10-952-459-18

Query Match 98.0%; Score 3939; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 3.7e-306;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 SRKTYTTLTYLKNTRYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTPDEF 72
Db 39 SRKTYTTLTYLKNTRYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTPDEF 98
Qy 73 GHSINDYSISDGGQFILLENNYVQWRHSYTSYDIYDLNKRQLITEERIINNQTQWTWS 132
Db 99 GHSINDYSISDGGQFILLENNYVQWRHSYTSYDIYDLNKRQLITEERIINNQTQWTWS 158
Qy 133 PVGHKLAVVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSAWMSP 192
Db 159 PVGHKLAVVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSAWMSP 218
Qy 193 NGTFLAYAQFNDTEVPLIEYSFYDESLOYPKTVRVPYPKAGAVNPTVKPFVNTDSLSS 252
Db 219 NGTFLAYAQFNDTEVPLIEYSFYDESLOYPKTVRVPYPKAGAVNPTVKPFVNTDSLSS 278
Qy 253 VTNATSIQITAPASMLIGDHYLCDVWATQERISLOWLRRIONYSVMDICDYDESSGRWN 312
Db 279 VTNATSIQITAPASMLIGDHYLCDVWATQERISLOWLRRIONYSVMDICDYDESSGRWN 338
Qy 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEBGYRHCYFQIDKKDCTFIT 372
Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEBGYRHCYFQIDKKDCTFIT 398
Qy 373 KGTWEVIGIEALTSYDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCISCELNPERCQYIS 432
Db 399 KGTWEVIGIEALTSYDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCISCELNPERCQYIS 458
Qy 433 VSPSKEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFII 492
Db 459 VSPSKEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFII 518
Qy 493 LNETKFWQMLPPHFDKSKYKYPILLDVAAGPCSKADTVFRLNWTATYLASTENIIVASF 552

```
Db      519 LNETKFWQMLPPHFDKSKKPLLLDDVYAGCSQKADTVFRLNWTYLASTENIIVASF 578
Qy      553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 612
Db      579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 638
Qy      613 VLGSGGVFKCGIAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKOV 672
Db      639 VLGSGGVFKCGIAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKOV 698
Qy      673 EYLLHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF 732
Db      699 EYLLHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF 758
Qy      733 IKQCFSLP 740
Db      759 IKQCFSLP 766

RESULT 5
US-10-631-467-565
; Sequence 565, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive P
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 565
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-467-565

Query Match      98.0%; Score 3939; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 3.7e-306;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      13 SRKTYTLTDYLNKTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSTFDEF 72
Db      39 SRKTYTLTDYLNKTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSTFDEF 98
Qy      73 GHSINDYSISPDGQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 132
Db      99 GHSINDYSISPDGQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 158
Qy      133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWYVEEYVFSAYSALWNSP 192
Db      159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWYVEEYVFSAYSALWNSP 218
Qy      193 NGTFLAYAQFNDTEVPLIEYSFYSDSLQYPKTVRPVYPKAGAVNPTVAFVVTNDSLS 252
Db      219 NGTFLAYAQFNDTEVPLIEYSFYSDSLQYPKTVRPVYPKAGAVNPTVAFVVTNDSLS 278
Qy      253 VTNATSIQITAPASMLIGHYLCDTVWATQERISLQWLRRIONYSVMDICDYDESSGRWN 312
Db      279 VTNATSIQITAPASMLIGHYLCDTVWATQERISLQWLRRIONYSVMDICDYDESSGRWN 338
Qy      313 CLVARQHIEMSTGWGRFRPSEPHEFDLGNPFYKLIISNEEGYRHICYFQIDKKCTFIT 372
Db      339 CLVARQHIEMSTGWGRFRPSEPHEFDLGNPFYKLIISNEEGYRHICYFQIDKKCTFIT 398
Qy      373 KGTWEVIGIEALTSYLYISNEYKMGPGRNLYKQLSDYTKVTCLSCELNPERCQYYS 432
```

```
Db      399 KGTWEVIGIEALTSYLYISNEYKMGPGRNLYKQLSDYTKVTCLSCELNPERCQYYS 458
Qy      433 VSPFSKEAKYQIQRCSGPGCLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTI 492
Db      459 VSPFSKEAKYQIQRCSGPGCLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTI 518
Qy      493 LNETKFWQMLPPHFDKSKKPLLLDDVYAGCSQKADTVFRLNWTYLASTENIIVASF 552
Db      519 LNETKFWQMLPPHFDKSKKPLLLDDVYAGCSQKADTVFRLNWTYLASTENIIVASF 578
Qy      553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 612
Db      579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 638
Qy      613 VLGSGGVFKCGIAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKOV 672
Db      639 VLGSGGVFKCGIAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKOV 698
Qy      673 EYLLHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF 732
Db      699 EYLLHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF 758
Qy      733 IKQCFSLP 740
Db      759 IKQCFSLP 766

RESULT 6
US-10-002-593-6
; Sequence 6, Application US/10002593
; Publication No. US20020137120A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTI
; FILE REFERENCE: Atty Docket No. US20020137120A1 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-593-6

Query Match      97.8%; Score 3933; DB 4; Length 766;
Best Local Similarity 99.9%; Pred. No. 1.1e-305;
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      13 SRKTYTLTDYLNKTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSTFDEF 72
Db      39 SRKTYTLTDYLNKTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSTFDEF 98
Qy      73 GHSINDYSISPDGQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 132
Db      99 GHSINDYSISPDGQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 158
Qy      133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWYVEEYVFSAYSALWNSP 192
Db      159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWYVEEYVFSAYSALWNSP 218
Qy      193 NGTFLAYAQFNDTEVPLIEYSFYSDSLQYPKTVRPVYPKAGAVNPTVAFVVTNDSLS 252
Db      219 NGTFLAYAQFNDTEVPLIEYSFYSDSLQYPKTVRPVYPKAGAVNPTVAFVVTNDSLS 278
Qy      253 VTNATSIQITAPASMLIGHYLCDTVWATQERISLQWLRRIONYSVMDICDYDESSGRWN 312
Db      279 VTNATSIQITAPASMLIGHYLCDTVWATQERISLQWLRRIONYSVMDICDYDESSGRWN 338
```

QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNBERGYRHICYFQIDKKDCTFIT 372
 DB 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNBERGYRHICYFQIDKKDCTFIT 398
 QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYVS 432
 DB 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYVS 458
 QY 433 VPSKEAKYYQLRCSGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQVPSKKLDFII 492
 DB 459 VPSKEAKYYQLRCSGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQVPSKKLDFII 518
 QY 493 LNETKFWYQMLPPLPHFDKSKKYPILLDDVYAGPCSKADTVPRLNWATYLASTENIIVASF 552
 DB 519 LNETKFWYQMLPPLPHFDKSKKYPILLDDVYAGPCSKADTVPRLNWATYLASTENIIVASF 578
 QY 553 DGRSGGYQGDKIMHAINRRLGTPEVEQIEAARQFSKMGFVDNKRKRIAIWGHSGYGYVTSM 612
 DB 579 DGRSGGYQGDKIMHAINRRLGTPEVEQIEAARQFSKMGFVDNKRKRIAIWGHSGYGYVTSM 638
 QY 613 VLGSAGVFKCGIAVAPVSRWEYDVSVYTERYMGSLPTPEDNLDHYRNSTVMSRAENFKQV 672
 DB 639 VLGSAGVFKCGIAVAPVSRWEYDVSVYTERYMGSLPTPEDNLDHYRNSTVMSRAENFKQV 698
 QY 673 EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF 732
 DB 699 EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF 758
 QY 733 IKQCFSLP 740
 DB 759 IKQCFSLP 766

RESULT 7

US-10-165-603-7
 ; Sequence 7, Application US/10165603
 ; Publication No. US20030021792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Roben, Paul W.
 ; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
 ; TITLE OF INVENTION: PROTEINS
 ; FILE REFERENCE: TPTECH.001A
 ; CURRENT APPLICATION NUMBER: US/10/165,603
 ; CURRENT FILING DATE: 2002-06-07
 ; PRIOR APPLICATION NUMBER: 60/297,021
 ; PRIOR FILING DATE: 2001-06-08
 ; PRIOR APPLICATION NUMBER: 60/305,117
 ; PRIOR FILING DATE: 2001-07-12
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-165-603-7

Query Match 97.8%; Score 3933; DB 4; Length 766;
 Best Local Similarity 99.9%; Pred. No. 1.1e-305;
 Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLTDVLYKNTYRLKLSLRWISDHEVLYKQENNILVFNABYGNSSVPLENSTDFEF 72
 DB 39 SRKTYTLTDVLYKNTYRLKLSLRWISDHEVLYKQENNILVFNABYGNSSVPLENSTDFEF 98
 QY 73 GHSINDYSISPDGQFILLLENNYKQWRHSYTSYDIYDLNKRQLITEERIPNNQTQWTS 132
 DB 99 GHSINDYSISPDGQFILLLENNYKQWRHSYTSYDIYDLNKRQLITEERIPNNQTQWTS 158
 QY 133 PVGHKLAYWNNNDIYVKEIPNLPSYRITWTGKEDIYNGITDWTVEEVSAYSALWSP 192
 DB 159 PVGHKLAYWNNNDIYVKEIPNLPSYRITWTGKEDIYNGITDWTVEEVSAYSALWSP 218

QY 193 NGTFLAYAQFNDETVPLIEYSFYSDESIQLYKTVRVPYPKAGAVNPTVFFVFNNTDSSLSS 252
 DB 219 NGTFLAYAQFNDETVPLIEYSFYSDESIQLYKTVRVPYPKAGAVNPTVFFVFNNTDSSLSS 278
 QY 253 VTNATSIQITAPASMLIGDHYLCDVTWATQBRISLOWLRRIONYSVMDICDYDESSGRWN 312
 DB 279 VTNATSIQITAPASMLIGDHYLCDVTWATQBRISLOWLRRIONYSVMDICDYDESSGRWN 338
 QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNBERGYRHICYFQIDKKDCTFIT 372
 DB 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNBERGYRHICYFQIDKKDCTFIT 398
 QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYVS 432
 DB 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYVS 458
 QY 433 VPSKEAKYYQLRCSGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQVPSKKLDFII 492
 DB 459 VPSKEAKYYQLRCSGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQVPSKKLDFII 518
 QY 493 LNETKFWYQMLPPLPHFDKSKKYPILLDDVYAGPCSKADTVPRLNWATYLASTENIIVASF 552
 DB 519 LNETKFWYQMLPPLPHFDKSKKYPILLDDVYAGPCSKADTVPRLNWATYLASTENIIVASF 578
 QY 553 DGRSGGYQGDKIMHAINRRLGTPEVEQIEAARQFSKMGFVDNKRKRIAIWGHSGYGYVTSM 612
 DB 579 DGRSGGYQGDKIMHAINRRLGTPEVEQIEAARQFSKMGFVDNKRKRIAIWGHSGYGYVTSM 638
 QY 613 VLGSAGVFKCGIAVAPVSRWEYDVSVYTERYMGSLPTPEDNLDHYRNSTVMSRAENFKQV 672
 DB 639 VLGSAGVFKCGIAVAPVSRWEYDVSVYTERYMGSLPTPEDNLDHYRNSTVMSRAENFKQV 698
 QY 673 EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF 732
 DB 699 EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF 758
 QY 733 IKQCFSLP 740
 DB 759 IKQCFSLP 766

RESULT 8

US-10-423-714-6
 ; Sequence 6, Application US/10423714
 ; Publication No. US20030180828A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Vanderbilt University
 ; APPLICANT: Brown, Nancy J.
 ; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTING ENZYME INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
 ; FILE REFERENCE: Acty Docket No. US20030180828A1 1242/48/2
 ; CURRENT APPLICATION NUMBER: US/10/423,714
 ; CURRENT FILING DATE: 2003-04-25
 ; PRIOR APPLICATION NUMBER: 60/244,524
 ; PRIOR FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-423-714-6

Query Match 97.8%; Score 3933; DB 4; Length 766;
 Best Local Similarity 99.9%; Pred. No. 1.1e-305;
 Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLTDVLYKNTYRLKLSLRWISDHEVLYKQENNILVFNABYGNSSVPLENSTDFEF 72
 DB 39 SRKTYTLTDVLYKNTYRLKLSLRWISDHEVLYKQENNILVFNABYGNSSVPLENSTDFEF 98
 QY 73 GHSINDYSISPDGQFILLLENNYKQWRHSYTSYDIYDLNKRQLITEERIPNNQTQWTS 132

Db 99 GHSINDYSIPDQFILLLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIIPNNTQWVTWS 158
 QY 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEEVSAYSALWSP 192
 Db 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEEVSAYSALWSP 218
 QY 193 NGTFLAYAQFNDTEVPLIEYSFYDESLOQPKTVRVPYKAGAVNPTVKFFVNTDLSLSS 252
 Db 219 NGTFLAYAQFNDTEVPLIEYSFYDESLOQPKTVRVPYKAGAVNPTVKFFVNTDLSLSS 278
 QY 253 VTNATSIQITAPASMLIGDHVLCVDTWATQERISLQWLRRIQNYSMVDICDYDESSGRWN 312
 Db 279 VTNATSIQITAPASMLIGDHVLCVDTWATQERISLQWLRRIQNYSMVDICDYDESSGRWN 338
 QY 313 CLVARQHIEMSTTGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 372
 Db 339 CLVARQHIEMSTTGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 398
 QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 432
 Db 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 458
 QY 433 VSFSKEAKYQYLRCSGPGGLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFII 492
 Db 459 VSFSKEAKYQYLRCSGPGGLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFII 518
 QY 493 LNETKFWQMIPLPHFDKSKYKPYLLDDVYAGCSOKADTVFRLNWTYLASTENIIVASF 552
 Db 519 LNETKFWQMIPLPHFDKSKYKPYLLDDVYAGCSOKADTVFRLNWTYLASTENIIVASF 578
 QY 553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSM 612
 Db 579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSM 638
 QY 613 VLGSYGVPKCGIAPVSRWEYYSVYTERVMGLPTPEDNLDHYRNSTVMSRAENFKQV 672
 Db 639 VLGSYGVPKCGIAPVSRWEYYSVYTERVMGLPTPEDNLDHYRNSTVMSRAENFKQV 698
 QY 673 EYLLHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 732
 Db 699 EYLLHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 758
 QY 733 IKQCFSLP 740
 Db 759 IKQCFSLP 766

RESULT 9
 US-10-295-027-922
 ; Sequence 922, Application US/10295027
 ; Publication No. US2003023250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Nataasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevezi, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15

; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347,211
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/347,349
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 922
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-295-027-922
 Query Match 97.8%; Score 3933; DB 4; Length 766;
 Best Local Similarity 99.9%; Pred. No. 1.1e-305;
 Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 13 SRKTYTLTDYLNKNTYRLKLYSLRWISDHELYLKQENNLVFNNAEYGNVSFLENSTFDFP 72
 Db 39 SRKTYTLTDYLNKNTYRLKLYSLRWISDHELYLKQENNLVFNNAEYGNVSFLENSTFDFP 98
 QY 73 GHSINDYSIPDQFILLLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIIPNNTQWVTWS 132
 Db 99 GHSINDYSIPDQFILLLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIIPNNTQWVTWS 158
 QY 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEEVSAYSALWSP 192
 Db 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEEVSAYSALWSP 218
 QY 193 NGTFLAYAQFNDTEVPLIEYSFYDESLOQPKTVRVPYKAGAVNPTVKFFVNTDLSLSS 252
 Db 219 NGTFLAYAQFNDTEVPLIEYSFYDESLOQPKTVRVPYKAGAVNPTVKFFVNTDLSLSS 278
 QY 253 VTNATSIQITAPASMLIGDHVLCVDTWATQERISLQWLRRIQNYSMVDICDYDESSGRWN 312
 Db 279 VTNATSIQITAPASMLIGDHVLCVDTWATQERISLQWLRRIQNYSMVDICDYDESSGRWN 338
 QY 313 CLVARQHIEMSTTGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 372
 Db 339 CLVARQHIEMSTTGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 398
 QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 432
 Db 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 458
 QY 433 VSFSKEAKYQYLRCSGPGGLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFII 492
 Db 459 VSFSKEAKYQYLRCSGPGGLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFII 518
 QY 493 LNETKFWQMIPLPHFDKSKYKPYLLDDVYAGCSOKADTVFRLNWTYLASTENIIVASF 552
 Db 519 LNETKFWQMIPLPHFDKSKYKPYLLDDVYAGCSOKADTVFRLNWTYLASTENIIVASF 578
 QY 553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSM 612
 Db 579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSM 638
 QY 613 VLGSYGVPKCGIAPVSRWEYYSVYTERVMGLPTPEDNLDHYRNSTVMSRAENFKQV 672
 Db 639 VLGSYGVPKCGIAPVSRWEYYSVYTERVMGLPTPEDNLDHYRNSTVMSRAENFKQV 698
 QY 673 EYLLHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 732

Db 699 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAMWYTDDEHGIIASSTAHOHIYTHMSHP 758

QY 733 IKQCFSLP 740

Db 759 IKQCFSLP 766

RESULT 10

US-10-794-899-41
; Sequence 41, Application US/10794899
; Publication No. US20040146516A1
; GENERAL INFORMATION:
; APPLICANT: Utah Ventures
; TITLE OF INVENTION: Lumen-Exposed Molecules and Methods Targeted Delivery
; FILE REFERENCE: 27110-715
; CURRENT APPLICATION NUMBER: US/10/794,899
; CURRENT FILING DATE: 2004-03-05
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-794-899-41

Query Match 97.8%; Score 3933; DB 4; Length 766;
Best Local Similarity 99.9%; Pred. No. 1.1e-305;
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABEYGNSSVFLNSTPDEF 72

Db 39 SRKTYTLDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABEYGNSSVFLNSTPDEF 98

QY 73 GHSINDYSISPDGQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWVTWS 132

Db 99 GHSINDYSISPDGQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWVTWS 158

QY 133 PVGHKLAYVWNNDIYVVKIEPNLPSYRITWTGKEDIYNGITDWVYEEVFSAYSALWNSP 192

Db 159 PVGHKLAYVWNNDIYVVKIEPNLPSYRITWTGKEDIYNGITDWVYEEVFSAYSALWNSP 218

QY 193 NGTFLAYAQFNDTEVPLIEYSFYSDLSLOYPKTVRVPYPKAGANNPTKPFVNTDLSLS 252

Db 219 NGTFLAYAQFNDTEVPLIEYSFYSDLSLOYPKTVRVPYPKAGANNPTKPFVNTDLSLS 278

QY 253 VTNATSIQITAPASMLIGDHYLCDVWTATQBRISLOWLRRIONYSVMDICDYDESSGRWN 312

Db 279 VTNATSIQITAPASMLIGDHYLCDVWTATQBRISLOWLRRIONYSVMDICDYDESSGRWN 338

QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKXDCFTIT 372

Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKXDCFTIT 398

QY 373 KGTWEVIGIEALTSDYLYIISNEYKMGPGGNLYKIQIDYTKVTCLSCELNPERCQYYS 432

Db 399 KGTWEVIGIEALTSDYLYIISNEYKMGPGGNLYKIQIDYTKVTCLSCELNPERCQYYS 458

QY 433 VPSFKEAKYYQLRCSGGLPLYTLHSSVNDKGLVLEDNSALDKMLQNVQMPKSKLDPII 492

Db 459 VPSFKEAKYYQLRCSGGLPLYTLHSSVNDKGLVLEDNSALDKMLQNVQMPKSKLDPII 518

QY 493 LNETKFWYQMLPPHFDKSKKYPILLDYYAGPCSQKADTVFRLNWTATYLASTENIIVASF 552

Db 519 LNETKFWYQMLPPHFDKSKKYPILLDYYAGPCSQKADTVFRLNWTATYLASTENIIVASF 578

QY 553 DGRSGGQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGNYSYGYVTSM 612

Db 579 DGRSGGQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGNYSYGYVTSM 638

QY 613 VLGSSSGVFKGIAVAPVSRWEYDYSVYTERYMGILPTPEDNLHYRNSVTMSRAENKQV 672

Db 639 VLGSSSGVFKGIAVAPVSRWEYDYSVYTERYMGILPTPEDNLHYRNSVTMSRAENKQV 698

QY 673 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAMWYTDDEHGIIASSTAHOHIYTHMSHP 732

Db 699 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAMWYTDDEHGIIASSTAHOHIYTHMSHP 758

QY 733 IKQCFSLP 740

Db 759 IKQCFSLP 766

RESULT 11

US-11-041-674-6
; Sequence 6, Application US/11041674
; Publication No. US20050181468A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERT
; FILE REFERENCE: Atty Docket No. 1242/48/2/2/2
; CURRENT APPLICATION NUMBER: US/11/041,674
; CURRENT FILING DATE: 2005-01-24
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 10/002,593
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 10/423,714
; PRIOR FILING DATE: 2003-4-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-041-674-6

Query Match 97.8%; Score 3933; DB 6; Length 766;
Best Local Similarity 99.9%; Pred. No. 1.1e-305;
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABEYGNSSVFLNSTPDEF 72

Db 39 SRKTYTLDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABEYGNSSVFLNSTPDEF 98

QY 73 GHSINDYSISPDGQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWVTWS 132

Db 99 GHSINDYSISPDGQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWVTWS 158

QY 133 PVGHKLAYVWNNDIYVVKIEPNLPSYRITWTGKEDIYNGITDWVYEEVFSAYSALWNSP 192

Db 159 PVGHKLAYVWNNDIYVVKIEPNLPSYRITWTGKEDIYNGITDWVYEEVFSAYSALWNSP 218

QY 193 NGTFLAYAQFNDTEVPLIEYSFYSDLSLOYPKTVRVPYPKAGANNPTKPFVNTDLSLS 252

Db 219 NGTFLAYAQFNDTEVPLIEYSFYSDLSLOYPKTVRVPYPKAGANNPTKPFVNTDLSLS 278

QY 253 VTNATSIQITAPASMLIGDHYLCDVWTATQBRISLOWLRRIONYSVMDICDYDESSGRWN 312

Db 279 VTNATSIQITAPASMLIGDHYLCDVWTATQBRISLOWLRRIONYSVMDICDYDESSGRWN 338

QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKXDCFTIT 372

Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKXDCFTIT 398

QY 373 KGTWEVIGIEALTSDYLYIISNEYKMGPGGNLYKIQIDYTKVTCLSCELNPERCQYYS 432

Db 399 KGTWEVIGIEALTSDYLYIISNEYKMGPGGNLYKIQIDYTKVTCLSCELNPERCQYYS 458

QY 433 VPSFKEAKYYQLRCSGGLPLYTLHSSVNDKGLVLEDNSALDKMLQNVQMPKSKLDPII 492

Db 459 VPSFKEAKYYQLRCSGGLPLYTLHSSVNDKGLVLEDNSALDKMLQNVQMPKSKLDPII 518

QY 493 LNETKFWYQMLPPHFDKSKKYPILLDYYAGPCSQKADTVFRLNWTATYLASTENIIVASF 552

Db 519 LNETHFWYQMLPPHFDKSKYPLLLDVVYAGPCSKADTVFRLNWTATYLASTENIIVASF 578
QY 553 DGRSGSYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSM 612
Db 579 DGRSGSYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSM 638
QY 613 VLGSGSVFKGCIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSVTMSRAENFKOV 672
Db 639 VLGSGSVFKGCIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSVTMSRAENFKOV 698
QY 673 EYLLIHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 732
Db 699 EYLLIHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 758
QY 733 IKQCFSLP 740
Db 759 IKQCFSLP 766

RESULT 12

US-10-476-264-147
; Sequence 147, Application US/10476264
; Publication No. US20050123910A1
; GENERAL INFORMATION:
; APPLICANT: Cookeon, William Osmond Charles Michael
; APPLICANT: Noffat, Miriam Fleur
; APPLICANT: Allen, Maxine
; APPLICANT: Lench, Nick
; TITLE OF INVENTION: Enzyme and SNP marker for disease
; FILE REFERENCE: 16721-002US1
; CURRENT APPLICATION NUMBER: US/10/476,264
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PCT/GB02/01887
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: GB0110044.5
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: GB0110046.0
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: GB0124594.3
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: GB0124575.2
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 421
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 147
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-476-264-147

Query Match 97.7%; Score 3929; DB 5; Length 766;
Best Local Similarity 99.7%; Pred. No. 2.3e-305;
Matches 726; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLTDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNAYCNSVFLNSFTDFEF 72
Db 39 SRKTYTLTDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNAYCNSVFLNSFTDFEF 98
QY 73 GHSINDYSISPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNTQWVTS 132
Db 99 GHSINDYSISPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNTQWVTS 158
QY 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSP 192
Db 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSP 218
QY 193 NGTFLEYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLS 252
Db 219 NGTFLEYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLS 278
QY 253 VTNATSIQITAPASMLIGHYLCVDTWATQERISLQWLRRIQNYSYMDICDYDESSGRWN 312
Db 279 VTNATSIQITAPASMLIGHYLCVDTWATQERISLQWLRRIQNYSYMDICDYDESSGRWN 338

QY 313 CLVARQHLEMTSTGWGRFRPSEPHFTLDGNSFYKIISNEGYRHICYFQIDKDCCTFTT 372
Db 339 CLVARQHLEMTSTGWGRFRPSEPHFTLDGNSFYKIISNEGYRHICYFQIDKDCCTFTT 398
QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKQLSDYTKVTCLSCELNPERCQYYS 432
Db 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKQLSDYTKVTCLSCELNPERCQYYS 458
QY 433 VSFSKEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPFSKLDPLI 492
Db 459 VSFSKEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPFSKLDPLI 518
QY 493 LNETHFWYQMLPPHFDKSKYPLLLDVVYAGPCSKADTVFRLNWTATYLASTENIIVASF 552
Db 519 LNETHFWYQMLPPHFDKSKYPLLLDVVYAGPCSKADTVFRLNWTATYLASTENIIVASF 578
QY 553 DGRSGSYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSM 612
Db 579 DGRSGSYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSM 638
QY 613 VLGSGSVFKGCIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSVTMSRAENFKOV 672
Db 639 VLGSGSVFKGCIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSVTMSRAENFKOV 698
QY 673 EYLLIHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 732
Db 699 EYLLIHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 758
QY 733 IKQCFSLP 740
Db 759 IKQCFSLP 766

RESULT 13

US-10-770-712-1
; Sequence 1, Application US/10770712
; Publication No. US20050170333A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
; FILE REFERENCE: IMSC12.008A
; CURRENT APPLICATION NUMBER: US/10/770,712
; CURRENT FILING DATE: 2004-02-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-712-1

Query Match 97.7%; Score 3929; DB 5; Length 766;
Best Local Similarity 99.7%; Pred. No. 2.3e-305;
Matches 726; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLTDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNAYCNSVFLNSFTDFEF 72
Db 39 SRKTYTLTDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNAYCNSVFLNSFTDFEF 98
QY 73 GHSINDYSISPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNTQWVTS 132
Db 99 GHSINDYSISPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNTQWVTS 158
QY 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSP 192
Db 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSP 218
QY 193 NGTFLEYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLS 252
Db 219 NGTFLEYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLS 278
QY 253 VTNATSIQITAPASMLIGHYLCVDTWATQERISLQWLRRIQNYSYMDICDYDESSGRWN 312

Db 279 VTNATSIQITAPASMLIGDHYLCVWTWATQBRISLQWLRRIQNYSMVMDICDYDESSGRWN 338
Qy 313 CLVARQHIEMSTTGWGFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFIT 372
Db 339 CLVARQHIEMSTTGWGFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFIT 398
Qy 373 KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYS 432
Db 399 KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYS 458
Qy 433 VSPFSKEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDPII 492
Db 459 VSPFSKEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDPII 518
Qy 493 LNETFWYQMLIPPHFDKSKYKYPILLDYYAGPCSKADTVFRLNWTATYLASTENIIVASF 552
Db 519 LNETFWYQMLIPPHFDKSKYKYPILLDYYAGPCSKADTVFRLNWTATYLASTENIIVASF 578
Qy 553 DGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 612
Db 579 DGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 638
Qy 613 VLGSAGYVFKGCGIAVAPVSRWEYDVSVYTERYMGGLPTPEDNLDHYRNSTVMSRAENFKQV 672
Db 639 VLGSAGYVFKGCGIAVAPVSRWEYDVSVYTERYMGGLPTPEDNLDHYRNSTVMSRAENFKQV 698
Qy 673 EYLLIHGTADDNVHFQQAQISKALVDVGVDFQAMWYTDHGHGASSTAHQHIYTHMSHF 732
Db 699 EYLLIHGTADDNVHFQQAQISKALVDVGVDFQAMWYTDHGHGASSTAHQHIYTHMSHF 758
Qy 733 IKQCFSLP 740
Db 759 IKQCFSLP 766

RESULT 14

US-10-770-712-2
; Sequence 2, Application US/10770712
; Publication No. US20050170333A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
; FILE REFERENCE: IMSCI2.008A
; CURRENT APPLICATION NUMBER: US/10770,712
; CURRENT FILING DATE: 2004-02-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-770-712-2

Query Match 84.7%; Score 3406.5; DB 5; Length 767;
Best Local Similarity 84.7%; Pred. No. 1.6e-263;
Matches 618; Conservative 52; Mismatches 57; Indels 3; Gaps 1;

Qy 13 SRKTYTLTYLKNTRYLKLKSLRWISDHEYLKQENNILVFNAYGNSGVPLENSTDEF 72
Db 37 SARTYTLADYLNKTRFKVKSLSRWVSDSEYLKQENNILVFNAYGNSGVPLENSTDEF 96
Qy 73 GHSINDYSISPDGQFILLLENNYKQWRHSYTSYDIYDLNKRQLITTEERIPNNTQWITS 132
Db 97 GDSISDYSVSPDRFLVLENNYKQWRHSYTSYDIYDLNKRQLITTEERIPNNTQWITS 156
Qy 133 PVGHKLAYVMNNDIYVYKIEPLPSYRITWTGKEDIYNGITDMYVEEVSAYGALWSP 192
Db 157 QEGHKLAYVMNNDIYVYKIEPLPSYRITWTGKEDIYNGITDMYVEEVSAYGALWSP 216
Qy 193 NGTFLAYAQFNDTVPVLEISFYSDLSLQPKTVVPYKAGAVNPTKPVVNTDLSLSS 252
Db 217 NGTFLAYAQFNDTVPVLEISFYSDLSLQPKTVVPYKAGAVNPTKPVVNTDLSLSS 276

Qy 253 VTNATSIQITAPASMLIGDHYLCVWTWATQBRISLQWLRRIQNYSMVMDICDYDESSGRWN 312
Db 277 TTTTIPMQITAPASVTTGSDHYLCVAVWVSEDRIISQWLRRIQNYSMVMDICDYDESSGRWN 336
Qy 313 CLVARQHIEMSTTGWGFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKD---CT 369
Db 337 CPTTREHETSATGWCGRFRPAPBPHFTSDGSSFYKIVSDKGKHCICQFQDKRPEQVCT 396
Qy 370 FITKGTWEVIGIEALTSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQ 429
Db 397 FITKGTWEVIGIEALTSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQ 456
Qy 430 YYSVSFSEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLD 489
Db 457 YYSVSFSEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLD 516
Qy 490 FIILNETFWYQMLIPPHFDKSKYKYPILLDYYAGPCSKADTVFRLNWTATYLASTENIIV 549
Db 517 FIVLNETFWYQMLIPPHFDKSKYKYPILLDYYAGPCSKADTVFRLNWTATYLASTENIIV 576
Qy 550 ASPDGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYV 609
Db 577 ASPDGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYV 636
Qy 610 TSMVLGSGYVFKGCGIAVAPVSRWEYDVSVYTERYMGGLPTPEDNLDHYRNSTVMSRAENF 669
Db 637 TSMVLGSGYVFKGCGIAVAPVSRWEYDVSVYTERYMGGLPTPEDNLDHYRNSTVMSRAENF 696
Qy 670 KQVEYLLIHGTADDNVHFQQAQISKALVDVGVDFQAMWYTDHGHGASSTAHQHIYTHM 729
Db 697 KQVEYLLIHGTADDNVHFQQAQISKALVDVGVDFQAMWYTDHGHGASSTAHQHIYTHM 756
Qy 730 SHFIKQCFSL 739
Db 757 SHFIKQCFSL 766

RESULT 15

US-10-165-603-4
; Sequence 4, Application US/10165603
; Publication No. US20030021792A1
; GENERAL INFORMATION:
; APPLICANT: Roben, Paul W.
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: TPTECH.001A
; CURRENT APPLICATION NUMBER: US/10/165,603
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/297,021
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/305,117
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-165-603-4

Query Match 84.5%; Score 3395.5; DB 4; Length 767;
Best Local Similarity 84.4%; Pred. No. 1.2e-262;
Matches 616; Conservative 53; Mismatches 58; Indels 3; Gaps 1;

Qy 13 SRKTYTLTYLKNTRYLKLKSLRWISDHEYLKQENNILVFNAYGNSGVPLENSTDEF 72
Db 37 SARTYTLADYLNKTRFKVKSLSRWVSDSEYLKQENNILVFNAYGNSGVPLENSTDEF 96
Qy 73 GHSINDYSISPDGQFILLLENNYKQWRHSYTSYDIYDLNKRQLITTEERIPNNTQWITS 132
Db 97 GDSISDYSVSPDRFLVLENNYKQWRHSYTSYDIYDLNKRQLITTEERIPNNTQWITS 156

```

Qy 133 PVCHKLAYVWNDIYVYKIEPNLPSYRIITWTGKEDIYNGITDWWYEEVEFSAYSALWWS 192
Db 157 QEGHKLAYVWNDIYVYKIEPHLPSHRITSTGKENVIFNGINDWVYEEBIFGAYSALWWS 216
Qy 193 NGTFLAYAQFNDETVPLIEYSFYSDLSQPKTVRPYPKAGAVNPTVKFFVWNTDSLSS 252
Db 217 NGTFLAYAQFNDETVPLIEYSFYSDLSQPKTVRPYPKAGAVNPTVKFFVWNTDSLSS 276
Qy 253 VTNATSIQITAPASMLIGHYLCVDTWATQERISQWLRRIQNYSVMDICDYDESSGRWN 312
Db 277 TTTTIPMQITAPASVTTGDBHLCDAWVSEDRIQLWLRRIQNYSVMAICDYDKTTLVWN 336
Qy 313 CLVARQHIEMSTGWRFRPPEPHFTLDGNSFYKIIISNEEGYRHICYFOIDKKD---CT 369
Db 337 CPTQEHETSATGMCGRFRPAPBPHFTSDGSSFYKIVSDKGYKHICQFKDKRKEQVCT 396
Qy 370 FITKGTWEVIGIEALTSYLYISNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERCQ 429
Db 397 FITKGAWEVISIEALTSYLYISNEYKMPGGRNLYKIQLTDHTNKKCLSCDLNPERCQ 456
Qy 430 YYSVSFSKAKYYQLRCSGGLPLTYTLHSSVNDKGLRVLEDNSALDKMLQNVOMPSKKLD 489
Db 457 YYSVLSLKEAKYYGLGCRGGLPLTYTLHRSTDQKELRVLEDNSALDKMLQDVOMPSKKLD 516
Qy 490 FIILNETKFWYQMLPPHEDKSKYPLLLDVYAGPCSKADTVPRLNWATYLASTENIIV 549
Db 517 FIVLNETRFYQMLPPHFDKSKYPLLLDVYAGPCSKADAAFRLNWATYLASTENIIV 576
Qy 550 ASFDGRSGYQGDKIMHAINRRLGTPEVEDQIEAARQFSKMGFVONKRIAIWGSYGGYV 609
Db 577 ASFDGRSGYQGDKIMHAINKRLGTLEVEDQIEAARQFLKMGFVDSKRVAIWGSYGGYV 636
Qy 610 TSMVLGSGGVKPCGIAVAPVSRWEYYDSVYTERYMGGLPTPEDNLDHYRNSTVMSRAENF 669
Db 637 TSMVLGSGGVKPCGIAVAPVSRWEYYDSVYTERYMGGLPTPEDNLDHYRNSTVMSRAENF 696
Qy 670 KQVEYLLIHGTADDNVHFQCSAQISKALVDVGVDFQAMWYTTDEDHGIASSTAHOHIYTHM 729
Db 697 KQVEYLLIHGTADDNVHFQCSAQISKALVDAGVDFQAMWYTTDEDHGIASSTAHOHIYSHM 756
Qy 730 SHFIKOCFSL 739
Db 757 SHFLQOCFSL 766

```

Search completed: February 15, 2006, 13:29:19
Job time : 171 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 13:14:53 ; Search time 48 Seconds
(without alignments)
1274.583 Million cell updates/sec

Title: US-10-659-055-3

Perfect score: 4020

Sequence: 1 ADPGSHHHHSHKRYTLT.....AQHIYTHMSHFHKQCFSLP 740

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA.*
1: /cgm2_6/ptodata/1/1aa/5 COMB.pdb.*
2: /cgm2_6/ptodata/1/1aa/6 COMB.pdb.*
3: /cgm2_6/ptodata/1/1aa/H COMB.pdb.*
4: /cgm2_6/ptodata/1/1aa/PCTUS COMB.pdb.*
5: /cgm2_6/ptodata/1/1aa/RE COMB.pdb.*
6: /cgm2_6/ptodata/1/1aa/backfiles.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3939	98.0	755	4	PCT-US93-07923-3
2	3939	98.0	759	4	PCT-US93-07923-2
3	3939	98.0	766	1	US-08-230-491A-3
4	3939	98.0	766	1	US-08-619-280A-3
5	3939	98.0	766	1	US-08-940-391-3
6	3939	98.0	766	2	US-09-794-236-1
7	3939	98.0	766	2	US-09-265-606-3
8	3939	98.0	775	2	US-09-949-016-10450
9	3933	97.8	766	2	US-10-002-593-6
10	3933	97.8	766	2	US-09-949-016-6146
11	3933	97.8	766	2	US-10-423-714-6
12	3929	97.7	766	2	US-09-518-550-27
13	3010	74.9	593	4	PCT-US93-07923-11
14	2168	53.9	760	1	US-08-230-491A-2
15	2168	53.9	760	1	US-08-619-280A-2
16	2168	53.9	760	1	US-08-940-391-2
17	2168	53.9	760	2	US-09-265-606-2
18	2158.5	53.7	761	2	US-09-518-550-26
19	1259.5	31.3	547	2	US-09-949-016-8330
20	1223	30.4	796	2	US-09-976-674-5
21	1198	29.8	743	2	US-10-363-937-4
22	1196	29.8	706	2	US-09-976-674-41
23	1152.5	28.7	691	2	US-09-976-674-43
24	987	24.6	771	2	US-09-079-592-2
25	963.5	24.0	771	2	US-09-462-284-2
26	914	22.7	818	2	US-09-462-845-3
27	914	22.7	818	2	US-10-402-312-3

28	914	22.7	818	2	US-10-401-437-3	Sequence 3, Appli
29	914	22.7	818	2	US-10-402-067-3	Sequence 3, Appli
30	914	22.7	818	2	US-10-401-436-3	Sequence 3, Appli
31	809	20.1	676	2	US-09-518-550-42	Sequence 42, Appl
32	809	20.1	723	2	US-09-518-550-29	Sequence 29, Appl
33	799.5	19.9	710	2	US-09-518-550-28	Sequence 28, Appl
34	733	18.2	931	2	US-09-079-592-11	Sequence 11, Appl
35	591	14.7	323	2	US-09-270-767-45296	Sequence 45296, A
36	584.5	14.5	494	2	US-09-248-796A-19472	Sequence 19472, A
37	528.5	13.1	882	2	US-09-976-674-1	Sequence 1, Appli
38	528.5	13.1	882	2	US-10-070-464-1	Sequence 1, Appli
39	500	12.4	879	2	US-09-976-674-33	Sequence 33, Appl
40	500	12.4	879	2	US-09-976-674-35	Sequence 35, Appl
41	497	12.4	732	2	US-09-518-550-30	Sequence 30, Appl
42	493.5	12.3	526	2	US-09-248-796A-19924	Sequence 19924, A
43	493.5	12.3	863	2	US-09-976-674-3	Sequence 3, Appli
44	493.5	12.3	892	2	US-09-976-674-23	Sequence 23, Appl
45	493.5	12.3	892	2	US-09-976-674-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
PCT-US93-07923-3
; Sequence 3, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schloeman, Stuart F.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 755
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US93-07923-3

Query Match 98.0%; Score 3939; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 13 SRKTYTLDTLKNYVRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVFLNSTFDEF 72
Db 28 SRKTYTLDTLKNYVRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVFLNSTFDEF 87
QY 73 GHSINDYSIPDQFILLBYNKKWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 132
Db 88 GHSINDYSIPDQFILLBYNKKWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 147
QY 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 192
Db 148 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 207
QY 193 NGTFELAYAFNDTEVPLIEYSFVSDSLQYPKTVRPYKAGAVNPTVAFVNTDLSLS 252
Db 208 NGTFELAYAFNDTEVPLIEYSFVSDSLQYPKTVRPYKAGAVNPTVAFVNTDLSLS 267
QY 253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 312
Db 268 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 327
QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKCKCTFIT 372
Db 328 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKCKCTFIT 387
QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 432
Db 388 KGTWEVIGIEALTSYLYIISNEYKMGPGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 447
QY 433 VFSFKEAKYQLRCSGPGPLVYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFII 492
Db 448 VFSFKEAKYQLRCSGPGPLVYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFII 507
QY 493 LNETKFWYQMLPPHFDKSKYPLLDVYAGPSCQKADTVFRLNWTATYLASTENIIVASF 552
Db 508 LNETKFWYQMLPPHFDKSKYPLLDVYAGPSCQKADTVFRLNWTATYLASTENIIVASF 567
QY 553 DGRSGYQGDQKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSM 612
Db 568 DGRSGYQGDQKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSM 627
QY 613 VLGSGGVFKCGIAPVSRWEYDSYVYTERYNGLTPTPDNDLDHYRNSTVMSRAENFKQV 672
Db 628 VLGSGGVFKCGIAPVSRWEYDSYVYTERYNGLTPTPDNDLDHYRNSTVMSRAENFKQV 687
QY 673 EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTTDEDHGIASTAHQIYTHMSHF 732
Db 688 EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTTDEDHGIASTAHQIYTHMSHF 747
QY 733 IKQCFSLP 740
Db 748 IKQCFSLP 755

```

RESULT 2

```

PCT-US93-07923-2
; Sequence 2, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schloesman, Stuart F.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)

```

```

; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-8906
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 759
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; PCT-US93-07923-2

Query Match 98.0%; Score 3939; DB 4; Length 759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 SRKTYTLDTLKNYVRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVFLNSTFDEF 72
Db 32 SRKTYTLDTLKNYVRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVFLNSTFDEF 91
QY 73 GHSINDYSIPDQFILLBYNKKWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 132
Db 92 GHSINDYSIPDQFILLBYNKKWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 151
QY 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 192
Db 152 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 211
QY 193 NGTFELAYAFNDTEVPLIEYSFVSDSLQYPKTVRPYKAGAVNPTVAFVNTDLSLS 252
Db 212 NGTFELAYAFNDTEVPLIEYSFVSDSLQYPKTVRPYKAGAVNPTVAFVNTDLSLS 271
QY 253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 312
Db 272 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 331
QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKCKCTFIT 372
Db 332 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKCKCTFIT 391
QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 432
Db 392 KGTWEVIGIEALTSYLYIISNEYKMGPGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 451
QY 433 VFSFKEAKYQLRCSGPGPLVYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFII 492
Db 452 VFSFKEAKYQLRCSGPGPLVYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFII 511
QY 493 LNETKFWYQMLPPHFDKSKYPLLDVYAGPSCQKADTVFRLNWTATYLASTENIIVASF 552
Db 512 LNETKFWYQMLPPHFDKSKYPLLDVYAGPSCQKADTVFRLNWTATYLASTENIIVASF 571
QY 553 DGRSGYQGDQKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSM 612
Db 572 DGRSGYQGDQKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSM 631
QY 613 VLGSGGVFKCGIAPVSRWEYDSYVYTERYNGLTPTPDNDLDHYRNSTVMSRAENFKQV 672
Db 632 VLGSGGVFKCGIAPVSRWEYDSYVYTERYNGLTPTPDNDLDHYRNSTVMSRAENFKQV 691

```



```

; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-619-280A-3

Query Match      98.0%; Score 3939; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 SRKTYTLDTYLNKTYRLKLYLRWISDHEYLKQENNLVFNABYGNSSVFLNSTFDF 72
Db 39 SRKTYTLDTYLNKTYRLKLYLRWISDHEYLKQENNLVFNABYGNSSVFLNSTFDF 98
QY 73 GHSINDYSIPDQGFILLEYVYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQV 132
Db 99 GHSINDYSIPDQGFILLEYVYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQV 158
QY 133 PVGHKLAVYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWYVEEVFSALW 192
Db 159 PVGHKLAVYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWYVEEVFSALW 218
QY 193 NGTFLAYAQFNDETEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKFFV 252
Db 219 NGTFLAYAQFNDETEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKFFV 278
QY 253 VTNATSIQITAPASMLIGDHVLCVDTWATQERISIQWLRRIONYSVMDICD 312
Db 279 VTNATSIQITAPASMLIGDHVLCVDTWATQERISIQWLRRIONYSVMDICD 338
QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDK 372
Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDK 398
QY 373 KGTWEVIGIEALTSYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCEL 432
Db 399 KGTWEVIGIEALTSYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCEL 458
QY 433 VSFSKEAKYQLRCSGPGILLYTHSSVNDKGLRVLEDNSALDKMLQNVQMP 492
Db 459 VSFSKEAKYQLRCSGPGILLYTHSSVNDKGLRVLEDNSALDKMLQNVQMP 518
QY 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGCSQKADTVFRLNWTYLASTENI 552
Db 519 LNETKFWYQMLPPHFDKSKYPLLLDVYAGCSQKADTVFRLNWTYLASTENI 578
QY 553 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGS 612
Db 579 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGS 638
QY 613 VLGSGSGVFKCGIAPVSRWEYDVSYTERYMGFLPTPEDNLDHYRNSVTM 672
Db 639 VLGSGSGVFKCGIAPVSRWEYDVSYTERYMGFLPTPEDNLDHYRNSVTM 698
QY 673 EYLLHGTADDNVHFOQSAISKALVDGVDFQAMWYTDDEHGIASSTAHOIY 732
Db 699 EYLLHGTADDNVHFOQSAISKALVDGVDFQAMWYTDDEHGIASSTAHOIY 758
QY 733 IKQCFSLP 740
Db 759 IKQCFSLP 766
```

RESULT 5
US-08-940-391-3
Sequence 3, Application US/08940391
Patent No. 5965373
GENERAL INFORMATION:
APPLICANT: Zimmermann, Rainer; Park, John E.;
Kettig, Wolfgang; Old, Lloyd J.
TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION
PROTEIN ALPHA, AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

```

; ADDRESSER: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,391
; FILING DATE: 01-OCT-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5965373man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-940-391-3
```

```

Query Match      98.0%; Score 3939; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 SRKTYTLDTYLNKTYRLKLYLRWISDHEYLKQENNLVFNABYGNSSVFLNSTFDF 72
Db 39 SRKTYTLDTYLNKTYRLKLYLRWISDHEYLKQENNLVFNABYGNSSVFLNSTFDF 98
QY 73 GHSINDYSIPDQGFILLEYVYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQV 132
Db 99 GHSINDYSIPDQGFILLEYVYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQV 158
QY 133 PVGHKLAVYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWYVEEVFSALW 192
Db 159 PVGHKLAVYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWYVEEVFSALW 218
QY 193 NGTFLAYAQFNDETEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKFFV 252
Db 219 NGTFLAYAQFNDETEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKFFV 278
QY 253 VTNATSIQITAPASMLIGDHVLCVDTWATQERISIQWLRRIONYSVMDICD 312
Db 279 VTNATSIQITAPASMLIGDHVLCVDTWATQERISIQWLRRIONYSVMDICD 338
QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDK 372
Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDK 398
QY 373 KGTWEVIGIEALTSYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCEL 432
Db 399 KGTWEVIGIEALTSYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCEL 458
QY 433 VSFSKEAKYQLRCSGPGILLYTHSSVNDKGLRVLEDNSALDKMLQNVQMP 492
Db 459 VSFSKEAKYQLRCSGPGILLYTHSSVNDKGLRVLEDNSALDKMLQNVQMP 518
QY 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGCSQKADTVFRLNWTYLASTENI 552
Db 519 LNETKFWYQMLPPHFDKSKYPLLLDVYAGCSQKADTVFRLNWTYLASTENI 578
QY 553 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGS 612
Db 579 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGS 638
QY 613 VLGSGSGVFKCGIAPVSRWEYDVSYTERYMGFLPTPEDNLDHYRNSVTM 672
Db 639 VLGSGSGVFKCGIAPVSRWEYDVSYTERYMGFLPTPEDNLDHYRNSVTM 698
QY 673 EYLLHGTADDNVHFOQSAISKALVDGVDFQAMWYTDDEHGIASSTAHOIY 732
Db 699 EYLLHGTADDNVHFOQSAISKALVDGVDFQAMWYTDDEHGIASSTAHOIY 758
QY 733 IKQCFSLP 740
Db 759 IKQCFSLP 766
```

QY 553 DGRSGYQDCKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGNSYGGVVTSM 612
 DB 579 DGRSGYQDCKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGNSYGGVVTSM 638
 QY 613 VLGSAGYQDCKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGNSYGGVVTSM 672
 DB 639 VLGSAGYQDCKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGNSYGGVVTSM 698
 QY 673 EYLLHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHF 732
 DB 699 EYLLHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHF 758
 QY 733 IKQCFSLP 740
 DB 759 IKQCFSLP 766
 RESULT 6
 US-09-794-236-1
 ; Sequence 1, Application US/09794236
 ; Patent No. 6337069
 ; GENERAL INFORMATION:
 ; APPLICANT: Grouzmann, Eric
 ; APPLICANT: Lacroix, Jean-Silvain
 ; APPLICANT: Monod, Michel
 ; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
 ; FILE REFERENCE: 81985/276823
 ; CURRENT APPLICATION NUMBER: US/09/794,236
 ; CURRENT FILING DATE: 2001-02-28
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-794-236-1

Query Match 98.0%; Score 3939; DB 2; Length 766;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 13 SRKTYTLDYKNTYRLKLSLRWISDHEYLKQENNILVFNAYGNSVFLNSTDFEF 72
 DB 39 SRKTYTLDYKNTYRLKLSLRWISDHEYLKQENNILVFNAYGNSVFLNSTDFEF 98
 QY 73 GHSINDYSISPDGQFILLVYVQWRHSYTSYDIDYLNKRLITEERIPNNTQWVTWS 132
 DB 99 GHSINDYSISPDGQFILLVYVQWRHSYTSYDIDYLNKRLITEERIPNNTQWVTWS 158
 QY 133 PVGHKLAVWNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWWYEEVFSAYGNSVFLNSTDFEF 192
 DB 159 PVGHKLAVWNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWWYEEVFSAYGNSVFLNSTDFEF 218
 QY 193 NGTFLAYAFNDTEVPLIEYSFDESQYKPTVRVPYKAGAVNPTKFFVNTDLSLS 252
 DB 219 NGTFLAYAFNDTEVPLIEYSFDESQYKPTVRVPYKAGAVNPTKFFVNTDLSLS 278
 QY 253 VTNATSIQITAPASMLIGDHYLCVDTWATQBRISLQWLRRIQNTSVMDICDYDSSGRWN 312
 DB 279 VTNATSIQITAPASMLIGDHYLCVDTWATQBRISLQWLRRIQNTSVMDICDYDSSGRWN 338
 QY 313 CLVARQHEIMSTTGVGRFSEPHFTLDGNSFYKIIISNREGYRHCYFQIDKDCFTIT 372
 DB 339 CLVARQHEIMSTTGVGRFSEPHFTLDGNSFYKIIISNREGYRHCYFQIDKDCFTIT 398
 QY 373 KGTWEVIGIEALTSIDLYIISNEYKMGPGGNNLYKIQLSDYTKVTKLSCNLPKRCQYYS 432
 DB 399 KGTWEVIGIEALTSIDLYIISNEYKMGPGGNNLYKIQLSDYTKVTKLSCNLPKRCQYYS 458
 QY 433 VSPFSEAKYQVLRCSGPGCLPLYTLHSSVNDKGLVLEDSALDQQLQNVQMPKSKLDFII 492
 DB 459 VSPFSEAKYQVLRCSGPGCLPLYTLHSSVNDKGLVLEDSALDQQLQNVQMPKSKLDFII 518

QY 493 LNETKFWYQMLPPHFDKSKYPLLLDYYVAGPCSKADTVPRLNWATVLASTENIIVASF 552
 DB 519 LNETKFWYQMLPPHFDKSKYPLLLDYYVAGPCSKADTVPRLNWATVLASTENIIVASF 578
 QY 553 DGRSGYQDCKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGNSYGGVVTSM 612
 DB 579 DGRSGYQDCKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGNSYGGVVTSM 638
 QY 613 VLGSAGYQDCKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGNSYGGVVTSM 672
 DB 639 VLGSAGYQDCKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGNSYGGVVTSM 698
 QY 673 EYLLHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHF 732
 DB 699 EYLLHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHF 758
 QY 733 IKQCFSLP 740
 DB 759 IKQCFSLP 766

RESULT 7
 US-09-265-606-3
 ; Sequence 3, Application US/09265606
 ; Patent No. 6846910
 ; GENERAL INFORMATION:
 ; APPLICANT: Zimmermann, Rainer; Park, John E.;
 ; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
 ; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
 ; TITLE OF INVENTION: ALPHA, AND USES THEREOF
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Felfe & Lynch
 ; STREET: 805 Third Avenue
 ; CITY: New York City
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10022
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/265,606
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/619,280
 ; FILING DATE: 18-MARCH-1996
 ; APPLICATION NUMBER: 08/230,491
 ; FILING DATE: 20-APRIL-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hanson, No. 6846910man D.
 ; REGISTRATION NUMBER: 30,946
 ; REFERENCE/DOCKET NUMBER: LUD 5330.1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 688-9200
 ; TELEFAX: (212) 838-3884
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 766 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 US-09-265-606-3

Query Match 98.0%; Score 3939; DB 2; Length 766;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 13 SRKTYTLDYKNTYRLKLSLRWISDHEYLKQENNILVFNAYGNSVFLNSTDFEF 72

Db 39 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNABYGNSSVFLNSTDDEF 98
Qy 73 GHSINDYSISPDGQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTWS 132
Db 99 GHSINDYSISPDGQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTWS 158
Qy 133 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWWS 192
Db 159 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWWS 218
Qy 193 NGTFLAYAQFNDTEVPLLEYSPYSDLSQYPTKTVRPYKAGAVNPTVKFFVNTDLS 252
Db 219 NGTFLAYAQFNDTEVPLLEYSPYSDLSQYPTKTVRPYKAGAVNPTVKFFVNTDLS 278
Qy 253 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 312
Db 279 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 338
Qy 313 CLVARQHITEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 372
Db 339 CLVARQHITEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 398
Qy 373 KGTWEVIGIEALTSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 432
Db 399 KGTWEVIGIEALTSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 458
Qy 433 VSFSKEAKYQYLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTII 492
Db 459 VSFSKEAKYQYLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTII 518
Qy 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGCSOKADTVFRLNWTYLASTENIIVASF 552
Db 519 LNETKFWYQMLPPHFDKSKYPLLLDVYAGCSOKADTVFRLNWTYLASTENIIVASF 578
Qy 553 DGRSGYQGDKIIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGYVTSM 612
Db 579 DGRSGYQGDKIIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGYVTSM 638
Qy 613 VLGSQGVFKCGIAPVSRWEYDSVYTERYMGVLPEDNLDHVRNSTVMSRAENFKQV 672
Db 639 VLGSQGVFKCGIAPVSRWEYDSVYTERYMGVLPEDNLDHVRNSTVMSRAENFKQV 698
Qy 673 EYLLIHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSSTAHOIYTHMSHF 732
Db 699 EYLLIHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSSTAHOIYTHMSHF 758
Qy 733 IKQCFSLP 740
Db 759 IKQCFSLP 766
RESULT 8
US-09-949-016-10450
; Sequence 10450, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10450
; LENGTH: 775
; TYPE: PRT

; ORGANISM: Human
US-09-949-016-10450
Query Match 98.0%; Score 3939; DB 2; Length 775;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 13 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNABYGNSSVFLNSTDDEF 72
Db 48 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNABYGNSSVFLNSTDDEF 107
Qy 73 GHSINDYSISPDGQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTWS 132
Db 108 GHSINDYSISPDGQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTWS 167
Qy 133 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWWS 192
Db 168 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWWS 227
Qy 193 NGTFLAYAQFNDTEVPLLEYSPYSDLSQYPTKTVRPYKAGAVNPTVKFFVNTDLS 252
Db 228 NGTFLAYAQFNDTEVPLLEYSPYSDLSQYPTKTVRPYKAGAVNPTVKFFVNTDLS 287
Qy 253 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 312
Db 288 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 347
Qy 313 CLVARQHITEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 372
Db 348 CLVARQHITEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 407
Qy 373 KGTWEVIGIEALTSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 432
Db 408 KGTWEVIGIEALTSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 467
Qy 433 VSFSKEAKYQYLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTII 492
Db 468 VSFSKEAKYQYLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTII 527
Qy 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGCSOKADTVFRLNWTYLASTENIIVASF 552
Db 528 LNETKFWYQMLPPHFDKSKYPLLLDVYAGCSOKADTVFRLNWTYLASTENIIVASF 587
Qy 553 DGRSGYQGDKIIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGYVTSM 612
Db 588 DGRSGYQGDKIIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGYVTSM 647
Qy 613 VLGSQGVFKCGIAPVSRWEYDSVYTERYMGVLPEDNLDHVRNSTVMSRAENFKQV 672
Db 648 VLGSQGVFKCGIAPVSRWEYDSVYTERYMGVLPEDNLDHVRNSTVMSRAENFKQV 707
Qy 673 EYLLIHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSSTAHOIYTHMSHF 732
Db 708 EYLLIHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSSTAHOIYTHMSHF 767
Qy 733 IKQCFSLP 740
Db 768 IKQCFSLP 775
RESULT 9
US-10-002-593-6
; Sequence 6, Application US/10002593
; Patent No. 6586198
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOGENESIS CONVERTING
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Atty Docket No. 6586198 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524

; PRIOR FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-002-593-6

Query Match 97.8%; Score 3933; DB 2; Length 766;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	13	SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNILVFNABEYGNSSVFLENSTPDEF	72
Db	39	SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNILVFNABEYGNSSVFLENSTPDEF	98
QY	73	GHSINDYSISPDGQFILLLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIIPNTQWVTWS	132
Db	99	GHSINDYSISPDGQFILLLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIIPNTQWVTWS	158
QY	133	PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWWSF	192
Db	159	PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWWSF	218
QY	193	NGTFLAYAQFNDTEVPLIEYSFYSDLSIQYPTKVRVYPKAGAVNPTVKFFVNTDSLSS	252
Db	219	NGTFLAYAQFNDTEVPLIEYSFYSDLSIQYPTKVRVYPKAGAVNPTVKFFVNTDSLSS	278
QY	253	VTNATSIQITAPASMLIGDHYLCVDTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN	312
Db	279	VTNATSIQITAPASMLIGDHYLCVDTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN	338
QY	313	CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKDCDTEIT	372
Db	339	CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKDCDTEIT	398
QY	373	KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	432
Db	399	KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	458
QY	433	VFSKEAKYYQLRCGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKCLDPII	492
Db	459	VFSKEAKYYQLRCGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKCLDPII	518
QY	493	LNETKFWYQMLPDPHFKSKYPLLLDVYAGCSQKADTVFRLNWTATYLASTENIIVASF	552
Db	519	LNETKFWYQMLPDPHFKSKYPLLLDVYAGCSQKADTVFRLNWTATYLASTENIIVASF	578
QY	553	DGRSGYQGDKIEMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGYVTSM	612
Db	579	DGRSGYQGDKIEMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGYVTSM	638
QY	613	VLGSGGVFKGIAVAPVSRWEYDYSVYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQV	672
Db	639	VLGSGGVFKGIAVAPVSRWEYDYSVYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQV	698
QY	673	EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF	732
Db	699	EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF	758
QY	733	IKQCFSLP 740	
Db	759	IKQCFSLP 766	

RESULT 10
 US-09-949-016-6146
 ; Sequence 6146, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6146
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-949-016-6146

Query Match 97.8%; Score 3933; DB 2; Length 766;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	13	SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNILVFNABEYGNSSVFLENSTPDEF	72
Db	39	SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNILVFNABEYGNSSVFLENSTPDEF	98
QY	73	GHSINDYSISPDGQFILLLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIIPNTQWVTWS	132
Db	99	GHSINDYSISPDGQFILLLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIIPNTQWVTWS	158
QY	133	PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWWSF	192
Db	159	PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWWSF	218
QY	193	NGTFLAYAQFNDTEVPLIEYSFYSDLSIQYPTKVRVYPKAGAVNPTVKFFVNTDSLSS	252
Db	219	NGTFLAYAQFNDTEVPLIEYSFYSDLSIQYPTKVRVYPKAGAVNPTVKFFVNTDSLSS	278
QY	253	VTNATSIQITAPASMLIGDHYLCVDTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN	312
Db	279	VTNATSIQITAPASMLIGDHYLCVDTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN	338
QY	313	CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKDCDTEIT	372
Db	339	CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKDCDTEIT	398
QY	373	KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	432
Db	399	KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	458
QY	433	VFSKEAKYYQLRCGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKCLDPII	492
Db	459	VFSKEAKYYQLRCGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKCLDPII	518
QY	493	LNETKFWYQMLPDPHFKSKYPLLLDVYAGCSQKADTVFRLNWTATYLASTENIIVASF	552
Db	519	LNETKFWYQMLPDPHFKSKYPLLLDVYAGCSQKADTVFRLNWTATYLASTENIIVASF	578
QY	553	DGRSGYQGDKIEMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGYVTSM	612
Db	579	DGRSGYQGDKIEMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGYVTSM	638
QY	613	VLGSGGVFKGIAVAPVSRWEYDYSVYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQV	672
Db	639	VLGSGGVFKGIAVAPVSRWEYDYSVYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQV	698
QY	673	EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF	732
Db	699	EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF	758
QY	733	IKQCFSLP 740	
Db	759	IKQCFSLP 766	

RESULT 11
US-10-423-714-6
; Sequence 6, Application US/10423714
; Patent No. 6887679
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Acty Docket No. 6887679 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/423,714
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-423-714-6

Query Match 97.8%; Score 3933; DB 2; Length 766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLDTYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDEF 72
DB 39 SRKTYTLDTYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDEF 98

QY 73 GHSINDYSIPDQGFILLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVTTWS 132
DB 99 GHSINDYSIPDQGFILLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVTTWS 158

QY 133 PVGHKLAVYNNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 192
DB 159 PVGHKLAVYNNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 218

QY 193 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLSL 252
DB 219 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLSL 278

QY 253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNYSDICDVEDSSGRWN 312
DB 279 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNYSDICDVEDSSGRWN 338

QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 372
DB 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 398

QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 432
DB 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 458

QY 433 VFSFKEAKYQYLRCSGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFII 492
DB 459 VFSFKEAKYQYLRCSGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFII 518

QY 493 LNETKFWYQMLPPHFDKSKYPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 552
DB 519 LNETKFWYQMLPPHFDKSKYPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 578

QY 553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIA1WGSYGGYVTSM 612
DB 579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIA1WGSYGGYVTSM 638

QY 613 VLGSQGVKFCGIAVAPSRWEYDYSYTERYNGLPTPDNLDHYRNSVTMGRANFKOV 672
DB 639 VLGSQGVKFCGIAVAPSRWEYDYSYTERYNGLPTPDNLDHYRNSVTMGRANFKOV 698

QY 673 EYLLIHGTADDNVHFQOSAISKALVDGVDFQAMWYTDDEHGIIASSTAHOIYTHMSHF 732

Db 699 EYLLIHGTADDNVHFQOSAISKALVDGVDFQAMWYTDDEHGIIASSTAHOIYTHMSHF 758

QY 733 IKQCFSLP 740
Db 759 IKQCFSLP 766

RESULT 12
US-09-518-550-27
; Sequence 27, Application US/09518550
; Patent No. 6875851
; GENERAL INFORMATION:
; APPLICANT: POTEMPA, James
; APPLICANT: BANBULA, Agnieszka
; TITLE OF INVENTION: PROLYL PEPTIDASES AND METHODS OF USE
; FILE REFERENCE: 235,00190101
; CURRENT APPLICATION NUMBER: US/09/518,550
; CURRENT FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/123,148
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: PCT/US00/05551
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 27
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-518-550-27

Query Match 97.7%; Score 3929; DB 2; Length 766;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 726; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLDTYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDEF 72
DB 39 SRKTYTLDTYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDEF 98

QY 73 GHSINDYSIPDQGFILLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVTTWS 132
DB 99 GHSINDYSIPDQGFILLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVTTWS 158

QY 133 PVGHKLAVYNNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 192
DB 159 PVGHKLAVYNNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 218

QY 193 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLSL 252
DB 219 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLSL 278

QY 253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNYSDICDVEDSSGRWN 312
DB 279 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNYSDICDVEDSSGRWN 338

QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 372
DB 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 398

QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 432
DB 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 458

QY 433 VFSFKEAKYQYLRCSGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFII 492
DB 459 VFSFKEAKYQYLRCSGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFII 518

QY 493 LNETKFWYQMLPPHFDKSKYPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 552
DB 519 LNETKFWYQMLPPHFDKSKYPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 578

QY 553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIA1WGSYGGYVTSM 612

Db 579 DGRSGYQDQIMHAINRLTGFVEDQIEAARQFSKMGFVDNKRRIAIWGSYGYVTSM 638
Qy 613 VLGSQGVFKGIAVAPVSRWEYDVSVTERYMGLPPTPEDNLDHYRNSTVMSRAENFKQV 672
Db 639 VLGSQGVFKGIAVAPVSRWEYVESVTERYMGLPPTPEDNLDHYRNSTVMSRAENFKQV 698
Qy 673 EYLLIHGTADNVHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHF 732
Db 699 EYLLIHGTADNVHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHF 758
Qy 733 IKQCFSLP 740
Db 759 IKQCFSLP 766

RESULT 13

PCT-US93-07923-11
; Sequence 11, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schlossman, Stuart F.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Frazer, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US93-07923-11

Query Match 74.9%; Score 3010; DB 4; Length 593;
Best Local Similarity 100.0%; Pred. No. 4.9e-256;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 13 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNLVFNABYSGNSVFLNSTDEF 72
Db 39 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNLVFNABYSGNSVFLNSTDEF 98
Qy 73 GHSINDYSISDPGIFILLEVYVQWRHSYTSYDIDLNKRQLITERIPNNTQWTVWS 132
Db 99 GHSINDYSISDPGIFILLEVYVQWRHSYTSYDIDLNKRQLITERIPNNTQWTVWS 158

Qy 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWMSF 192
Db 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWMSF 218
Qy 193 NGTFPLAYAQNDTEVPLIEYSFYSDESLOYPKTVRVPYPKAGAVNPTVKPVVNTDLSLSS 252
Db 219 NGTFPLAYAQNDTEVPLIEYSFYSDESLOYPKTVRVPYPKAGAVNPTVKPVVNTDLSLSS 278
Qy 253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSVMDICDYDESSGRWN 312
Db 279 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSVMDICDYDESSGRWN 338
Qy 313 CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCFTFIT 372
Db 339 CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCFTFIT 398
Qy 373 KGTWEVIGIEALTSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 432
Db 399 KGTWEVIGIEALTSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 458
Qy 433 VSPSKBKYQLRCGSGPLPLYTLHSSVNDKGLRVLEDNSALDMLQNVQMPSSKCLDFII 492
Db 459 VSPSKBKYQLRCGSGPLPLYTLHSSVNDKGLRVLEDNSALDMLQNVQMPSSKCLDFII 518
Qy 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 552
Db 519 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 578
Qy 553 DGRSGYQDQIMHA 567
Db 579 DGRSGYQDQIMHA 593

RESULT 14

US-08-230-491A-2
; Sequence 2, Application US/08230491A
; Patent No. 5587299
; GENERAL INFORMATION:
; APPLICANT: Rettig, Wolfgang J.; Scanlan, Matthew J.;
; APPLICANT: Garin-Chesa, Pilar; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR
; TITLE OF INVENTION: FIBROBLAST ACTIVATION PROTEIN ' AND USES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FELPE & LYNCH
; STREET: 805 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE 3.5 inch 1.2 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT - ASC II
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/230,491A
; FILING DATE: 20-APRIL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5587299man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 330
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

US-08-230-491A-2

Query Match 53.9%; Score 2168; DB 1; Length 760;
Best Local Similarity 52.5%; Pred. No. 9.9e-182;
Matches 389; Conservative 138; Mismatches 200; Indels 14; Gaps 7;
QY 3 PGSSHHHHHSHKRTVLTDLKNTYRLKLYSLRWISDHELYKQ-ENNILVFNAEYGNSS 61
Db 28 PSRVHNSBENTWALTLKDLINGTFSYKTFPFWNLSGGYELHQSDANNILVNIETGQSY 87
QY 62 VFLENSTDFEGHSIN--DYSISPDGQFILLIENYVVKQWRHSYTSYDIYDLNKRQLITE 119
Db 88 TILSNRTM---KSNVNASNYGLSPDRQFVYLESYKLRYSYATYIYDLSNGEFVRG 143
QY 120 ERIPNNTQWVTSVPGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEE 179
Db 144 NELPRPIQYLCWSPVSKLAYVYQNNIYVKRQPGDPPQITFNGRENKIFNGIPDWYEE 203
QY 180 EVFSAYSALWSPNGTFLAYAFNDTEVPLIEVSFYSDSLQYPKTVRPYKAGAVNPT 239
Db 204 EMLPTKYALWSPNGKFLAYAFNDKDIPIVAYSYGDE--QYPTINIPIYKAGAKNPV 261
QY 240 VKFVVNTDLSVSTNATSIOITAPASMLIGHYLCVDTWATQERISLQWLRRIQNSVM 299
Db 262 VRIFIIDTTYPAVGPQ---EVPVPAMIASDYFVSWLTWVTDERVQLWLRKVQNSVL 318
QY 300 DICVDDESSGRNCLVARQHIEMSTTGWVGRFRPSEPHEFTLDGNSFYKIIISNEEGYRHC 359
Db 319 SICDFREDQWDCPKTQEHIEESRTGWAGFFVSRPVFSYDAISYKIFSDKGKGIH 378
QY 360 YFQIDKDCOTFITKGTWVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDY-TKVT 418
Db 379 YIKDTVENAIQITSGKWEAINIFRVTDQSLFYSSNEFEYPPGRRIYRISIGSYPPSKC 438
QY 419 LSCELNPERCOYVSFSKSKAYYOLRCSGPLPLYTLHSSVNDKGLRVLEDNSALDKML 478
Db 439 VTCHLRKERCQYITASPSDYAKYALVYCGGPIPISTLHDGRTDQEIKEILENKLENAL 498
QY 479 QNVQPSKKLDPIILNETKFWYQMLPPHFDKSKYPLLLDYAGPCSKQADTVFRLNWA 538
Db 499 KNIQLPKEIEKKLEVEDEITLWYKMLPPQFDRSKYPLLIQYVGGPCSQSVRSVFAVNI 558
QY 539 TYLASTENIIVASPDGRSGYQGDKIMHAINRRLGTFTVEDOI EAARQFSKMGFVDNKR 598
Db 559 SYLASKEGVIALVDGRGTAFQGDKLLIYARKLGVIEVEDQITAVRKFIEMGFIDEKRI 618
QY 599 AIWGSYGGYVTSMLVSGSGVFKGCIAPVAPSRMEYDSYTERYMGLPPTPEDNLDHYR 658
Db 619 AIWGSYGGYVSSLASGTGLFKGCIAPVAPVSSWEYASVYTERFMGLPTKDDNLEHYK 678
QY 659 NSTVMSRAENPKQVEYLLHGTADNNVHPQQAQISKALVDVGVDFQAMWYTDDBHGAS 718
Db 679 NSTVMARAEYFRNVYLLHGTADNNVHPQNSAQIAKALVNAQVDFQAMWYSDQNHGL-S 737
QY 719 STAHOIYTHMSHFILKQCFSL 739
Db 738 GUSTNRLYTHMFLKQCFSL 758

RESULT 15
US-08-619-280A-2
Sequence 2, Application US/08619280A
Patent No. 5767242
GENERAL INFORMATION:
APPLICANT: Zimmermann, Rainer; Park, John E.;
APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
TITLE OF INVENTION: ALPHA, AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City

STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/619,280A
FILING DATE: 18-MARCH-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/230,491
FILING DATE: 20-APRIL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5767242man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5330.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 760 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-619-280A-2

Query Match 53.9%; Score 2168; DB 1; Length 760;
Best Local Similarity 52.5%; Pred. No. 9.9e-182;
Matches 389; Conservative 138; Mismatches 200; Indels 14; Gaps 7;
QY 3 PGSSHHHHHSHKRTVLTDLKNTYRLKLYSLRWISDHELYKQ-ENNILVFNAEYGNSS 61
Db 28 PSRVHNSBENTWALTLKDLINGTFSYKTFPFWNLSGGYELHQSDANNILVNIETGQSY 87
QY 62 VFLENSTDFEGHSIN--DYSISPDGQFILLIENYVVKQWRHSYTSYDIYDLNKRQLITE 119
Db 88 TILSNRTM---KSNVNASNYGLSPDRQFVYLESYKLRYSYATYIYDLSNGEFVRG 143
QY 120 ERIPNNTQWVTSVPGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEE 179
Db 144 NELPRPIQYLCWSPVSKLAYVYQNNIYVKRQPGDPPQITFNGRENKIFNGIPDWYEE 203
QY 180 EVFSAYSALWSPNGTFLAYAFNDTEVPLIEVSFYSDSLQYPKTVRPYKAGAVNPT 239
Db 204 EMLPTKYALWSPNGKFLAYAFNDKDIPIVAYSYGDE--QYPTINIPIYKAGAKNPV 261
QY 240 VKFVVNTDLSVSTNATSIOITAPASMLIGHYLCVDTWATQERISLQWLRRIQNSVM 299
Db 262 VRIFIIDTTYPAVGPQ---EVPVPAMIASDYFVSWLTWVTDERVQLWLRKVQNSVL 318
QY 300 DICVDDESSGRNCLVARQHIEMSTTGWVGRFRPSEPHEFTLDGNSFYKIIISNEEGYRHC 359
Db 319 SICDFREDQWDCPKTQEHIEESRTGWAGFFVSRPVFSYDAISYKIFSDKGKGIH 378
QY 360 YFQIDKDCOTFITKGTWVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDY-TKVT 418
Db 379 YIKDTVENAIQITSGKWEAINIFRVTDQSLFYSSNEFEYPPGRRIYRISIGSYPPSKC 438
QY 419 LSCELNPERCOYVSFSKSKAYYOLRCSGPLPLYTLHSSVNDKGLRVLEDNSALDKML 478
Db 439 VTCHLRKERCQYITASPSDYAKYALVYCGGPIPISTLHDGRTDQEIKEILENKLENAL 498
QY 479 QNVQPSKKLDPIILNETKFWYQMLPPHFDKSKYPLLLDYAGPCSKQADTVFRLNWA 538
Db 499 KNIQLPKEIEKKLEVEDEITLWYKMLPPQFDRSKYPLLIQYVGGPCSQSVRSVFAVNI 558
QY 539 TYLASTENIIVASPDGRSGYQGDKIMHAINRRLGTFTVEDOI EAARQFSKMGFVDNKR 598
Db 559 SYLASKEGVIALVDGRGTAFQGDKLLIYARKLGVIEVEDQITAVRKFIEMGFIDEKRI 618

THIS PAGE BLANK (USPTO)

Result No.	Query \$			DB	ID	Description
	Score	Match	Length			
1	2153.6	98.6	3411	4	CR860150	Pongo pyg
2	1677.6	76.8	3289	4	AK085370	Mus muscu
3	1549.4	70.9	2193	10	AY418368	Homo sapi
4	1536.8	70.4	2193	10	AY418369	Pan trogl
5	1237	56.6	2191	10	AY418370	Mus muscu
6	1041.2	47.7	1081	1	AL554569	AL554569
7	865	39.6	987	5	EX339854	EX339854
8	829.2	38.0	886	5	EX348426	EX348426
9	818	37.5	1022	1	AL548644	AL548644
10	803.2	36.8	1022	1	AL541949	AL541949
11	795.8	36.4	1007	1	AL544456	AL544456
12	768.4	35.2	846	1	AU122649	AU122649
13	722	33.1	900	5	BQ937942	BQ937942
14	721	33.0	897	1	AL553051	AL553051
15	712.6	32.6	791	1	AU139683	AU139683
16	704	32.2	731	6	CB957930	CB957930
17	701.4	32.1	741	5	KX496051	KX496051
18	695.8	31.9	904	5	BX397280	BX397280
19	683	31.3	773	1	AU139062	AU139062
20	682.6	31.3	2277	10	AY402291	AY402291
21	678.4	31.1	2277	10	AY402292	AY402292
22	675.6	30.9	887	1	AL550603	AL550603

RESULT 1	
CR860150	3411 bp mRNA linear HTC 12-NOV-2004
LOCUS	Pongo pygmaeus mRNA; cDNA DKFZp469P1419 (from clone DKFZp469P1419).
DEFINITION	
ACCESSION	CR860150
VERSION	CR860150.1 GI:55731156
KEYWORDS	HTC.
SOURCE	Pongo pygmaeus (orangutan)
ORGANISM	Pongo pygmaeus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE	1. (bases 1 to 3411) Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R., Mewes, H.-W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
AUTHORS	The German cDNA Consortium Direct Submission Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
CONSRMT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp469P1419) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469P1419 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/ .
COMMENT	Location/Qualifiers 1. 3411 /organism="Pongo pygmaeus" /mol_type="mRNA" /db_xref="taxon:9600" /clone="DKFZp469P1419" /tissue type="kidney" /clone lib="469 (synonym: pkidi). Vector pSport1_sfi; host DH10B; sites sfila + sfilb" /dev stage="adult" /note="dipeptidyl peptidase IV (Homo sapiens)"
gene	1. 3411 /gene="DKFZp469P1419"
CDS	63..2360 /gene="DKFZp469P1419"
FEATURES	
source	

```

/codon_start=1
/product="hypothetical protein"
/protein_id="CAH92293.1"
/db_xref="GI:55731157"
/translation="MKTPKVLGLGALAAVLIITVPVLLNKGNADATASRKYTYL
TDYLNKTYRLKLYLSRWISDHYLYKQBNILVFNABYCNBSAFLENSFPDFGHIN
DYSIGDGFILLEYVYKWRHSYVASDIYDLNKROLI TEDRI PNNTQWTVSPVG
HKLAVWNDIILVLEPNLPSHRIWTGKEDIIVNGIDTMYVEEVEFSAYALWSPN
GFLAYAFNDTEVEFLIYFSYDESLOQPKTVRPVPKAGAVNFTVPFVNTDSL
SVTNATSQITATASMLIGDHYLCDVTWATERISLQWLRRIQNSVMDICDYDESSG
RWNCILVQHIEMSTYGVGRFPSEPHFTSDGNSFYKIIINBEGYRHCYFQIDKDD
CTFYIKTGVEFGEIYALYSIDYLYISNEYKMGPNRNLKIQLSYTKVCLSCLENP
BRQYDYSFESKEAYOLRCSGPGLYPLTYLHSSVNDKPRVLEDNSLTDKMLQNM
PSKLDLFIILNETKMYQMLPPHFDKSKYPLLDVAVGPCSOKADPFRVLNWKRIA
ASTENKIIIVASFGRSGSGQDKIMHAIINRLGTFFEBDQIEAARQFSGMFPVNDKRIA
IWNISYIGGVYSMVJGSGGVFKGIAVAPVSRWBYIYDSVYTERTWGLPTPEDNLIDHY
RNSTYMSRAENPKQVEYLIILHGTADDNVHFQOOSAISKALVDAGVDFQAMMYTDEDDHG
IASSTAHQHIHYTHMSHFINKQFSLP"

```

ORIGIN

```

Query Match      98.6%; Score 2153.6; DB 4; Length 3411;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2165; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 AGTCGCAGAACTTACACTCTAACTGATTACTTAAAAAATCTATTAGACTGAAGTTATAC 60
DB 174 AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATCTATTAGACTGAAGTTATAC 233

QY 61 TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAACAAGAAAAATATATCTTGGTA 120
DB 234 TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAACAAGAAAAATATATCTTGGTA 293

QY 121 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTCATGATGATTT 180
DB 294 TTCAATGCTGAATATGGAACAGCTCAGCTTTCTTGGAGAACAGTACATTCATGATGATTT 353

QY 181 GGCATCTCTCAATGATTATTCATATCTCTGATGGGAGTTTATCTCTTAGAATAC 240
DB 354 GGCATCTCTCAATGATTATTCATATCTCTGATGGGAGTTTATCTCTTAGAATAC 413

QY 241 AACTAGCTGAAGCAATGGAGGCAATCTCTACACAGCTTCATATGACATTTATGATTTAAAT 300
DB 414 AACTATGTGAAGCAATGGAGGCAATCTCTACACAGCTTCATATGACATTTATGATTTAAAT 473

QY 301 AAAAGCAGCTGATTTACAGAGAGGATTTCCAAACACACACAGTGGGTTCATGGTCA 360
DB 474 AAAAGCAGCTGATTTACAGAGAGGATTTCCAAACACACACAGTGGGTTCATGGTCA 533

QY 361 CCAGTGGGTTCATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCA 420
DB 534 CCAGTGGGTTCATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCA 593

QY 421 AATTTACCAAGTTTACAGAAATCAATGGACGGGAAAGAGATATATATATATATGGAATA 480
DB 594 AATTTACCAAGTTTACAGAAATCAATGGACGGGAAAGAGATATATATATATGGAATA 653

QY 481 ACTGACTGGGTTTATGAAGAGGAGTCTTCAGTGCCTACTCTGCTCTGTTGGTCTCCA 540
DB 654 ACTGACTGGGTTTATGAAGAGGAGTCTTCAGTGCCTACTCTGCTCTGTTGGTCTCCA 713

QY 541 AACGGCAGCTTTTATGATATGCCCAATTTTAAAGACACAGAGAGTCCCACTTATTTGAATAC 600
DB 714 AACGGCAGCTTTTATGATATGCCCAATTTTAAAGACACAGAGAGTCCCACTTATTTGAATAC 773

QY 601 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTAGCGGTTCCCATATCCAAAG 660
DB 774 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTAGCGGTTCCCATATCCAAAG 833

QY 661 GAGGAGCTGTGATCCAACTGTAAAGTTCTTTGTTGTTAAATACAGACTCTCTCAGCTCA 720
DB 834 GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTTGTTAAATACAGACTCTCTCAGCTCA 893

QY 721 GTCACCAATGCAACTTCCATACAAATCACTGCTCCTCTTCTATGTTGATAGGGGATCAC 780

```

```

DB 894 GTCACCAATGCAACTTTCATACAAATCACTGCTCCTGCTTCTATGTTGATAGGGGATCAC 953
QY 781 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTGACGTGGCTCAGGAGG 840
DB 954 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTGACGTGGCTCAGGAGG 1013
QY 841 ATTCAAGAACTATTTCGGTTCATGGATATTGTGACATATGATGAATCCAGTGGAGAGATGGAAC 900
DB 1014 ATTCAAGAACTATTTCGGTTCATGGATATTGTGACATATGATGAATCCAGTGGAGAGATGGAAC 1073
QY 901 TGCTTAGTGGCACGGCAACACATTTGAAATAGTACTACTGCGTGGTGGTGGAGAGATTAGG 960
DB 1074 TGCTTAGTGGCACGGCAACACATTTGAAATAGTACTACTGCGTGGTGGTGGAGAGATTAGG 1133
QY 961 CTTTCAGAACCTCATTTTACCTTTCATGGTGAATAGTCTTCAAGACATCATCAGCAATGAA 1020
DB 1134 CTTTCAGAACCTCATTTTACCTTTCATGGTGAATAGTCTTCAAGACATCATCAGCAATGAA 1193
QY 1021 GAAGGTACAGACACATTTGCTATTTCCTCAAAATAGATAAAAAAGACTGCACATTTATTACA 1080
DB 1194 GAAGGTACAGACACATTTGCTATTTCCTCAAAATAGATAAAAAAGACTGCACATTTATTACA 1253
QY 1081 AAAGCACCTCGGGAAGTCAATCGGGATAGAAGCTCTTAAACAGTGAATATCTATACATTT 1140
DB 1254 AAAGCACCTCGGGAAGTCAATCGGGATAGAAGCTCTTAAACAGTGAATATCTATACATTT 1313
QY 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAAGGATCTTTATAAAATCCAACTTATTGAC 1200
DB 1314 AGTAATGAATATAAAGGAATGCCAGGAGGAAGGATCTTTATAAAATCCAACTTATTGAC 1373
QY 1201 TATACAAAGTGAACATGCTCAGTTGTGAGCTGAAATCCGGAAAGGTGTGAGTACTATTCT 1260
DB 1374 TATACAAAGTGAACATGCTCAGTTGTGAGCTGAAATCCGGAAAGGTGTGAGTACTATTCT 1433
QY 1261 GTGTCATTTCACTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCGGCTCGCCC 1320
DB 1434 GTGTCATTTCACTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCGGCTCGCCC 1493
QY 1321 CTCTATCTCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCGGAAGACAAATCA 1380
DB 1494 CTCTATCTCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCGGAAGACAAATCA 1553
QY 1381 GCTTTGGATATAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAACTTGACATTTATT 1440
DB 1554 GCTTTGGATATAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAACTTGACATTTATT 1613
QY 1441 TTGAAATGAAACAAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGAATAAATCCAAG 1500
DB 1614 TTGAAATGAAACAAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGAATAAATCCAAG 1673
QY 1501 AAATATCTCTACTATTATGATGTTGATGCGGCCCATGTAGTCAAAAAGCAGACACTGTC 1560
DB 1674 AAATATCTCTACTATTATGATGTTGATGCGGCCCATGTAGTCAAAAAGCAGACACTGTC 1733
QY 1561 TTCAGACTGAACCTGGGCCACTTTACCTTGCAGACACAGAAAAATTTATAGTAGTAGCTTT 1620
DB 1734 TTCAGACTGAACCTGGGCCACTTTACCTTGCAGACACAGAAAAATTTATAGTAGTAGCTTT 1793
QY 1621 GATGCGAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGACTG 1680
DB 1794 GATGCGAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGACTG 1853
QY 1681 GGAACTTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGCAATTTTCAAAAATGGGATTT 1740
DB 1854 GGAACTTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGCAATTTTCAAAAATGGGATTT 1913
QY 1741 GTGGACAAACAAAGAAATTTGGGCTGCTCATATGAGAGGTGACCTAACTCAATG 1800
DB 1914 GTGGACAAACAAAGAAATTTGGGCTGCTCATATGAGAGGTGACCTAACTCAATG 1973
QY 1801 GTCTCTGGGATCGGGAAGTGGCTGTTCAAGTGTGAAATAGCCGTGGCGCTGTATCCCCGG 1860

```


Db 132 AGCCGACGACGTAATTTCACTAGCTGACTATTTAAAGAGTACCTTTGGGTCAAGTCTTAC 191
 Qy 61 TCCTTAAGATGATTTTCAATCATGATCATGATATCTCTACAAACAGAAATAATATCTTGGTA 120
 Db 192 TCTTTTGGTGGGTTTCAGACTTTGAATACCTCTACAAACAGAGAAACAATATCTTGGCTG 251
 Qy 121 TTCAATGCTGAATATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
 Db 252 CTCAATGCTGAACATGGAACAGCTCCATTTCTTGGAGAACAGTACCTTTGGAAGCTTT 311
 Qy 181 GGACATTTCTATCAATGATTTTCAATATCTCTGATGGGCACTTTATTTCTTTAGAAATAC 240
 Db 312 GGA-----TATCATTTCACTGTCACCTGACCGACTGTTTGTCTCTTGGAAATAC 359
 Qy 241 AACTAGGTGAACATGGAAGGATTTCTTACAGAGCTTCTATGACATTTATGATTTAAAT 300
 Db 360 AACTAGGTGAACATGGAAGGATTTCTTACAGAGCTTCTATGACATTTATGATTTAAAT 419
 Qy 301 AAAAGGCACTGATTTACAGAGAGAGGATTTCCAAACACACAGCTGGGTACATGGTCA 360
 Db 420 AAAAGACAGCTGATTCAGAGAGAGAGATTTCCAAATATATACAGGTGGATACATGGTCA 479
 Qy 361 CCAGTGGGTCTATAAATTTGSCATATGTTTGGAAACAATGACATTTATTTAAATTTGAACCA 420
 Db 480 CCAGAGGTCTATAAGTTGGCATATGCTCTGGAAGATGATTTTACGTTAAAGTTGAACCA 539
 Qy 421 AATTTACCAAGTTACAGATTCATGAGGAGGATTTCCAAACACACAGCTGGGTACATGGTCA 480
 Db 540 CACTTACCTAGTCATAGGATCATCTGACAGGAGAGAAATGTAATATATATATGGAATA 599
 Qy 481 ACTGACTGGGTTTATGAGAGGAAGTCTTCAGTGCTCTCTGCTCTGTTGGTGGTCTCCA 540
 Db 600 ACTGACTGGGTTTATGAGAGGAAGTCTTCGGTGGCTCTCTGCACTGTTGGTGGTCTCCA 659
 Qy 541 AACGGCACTTTTATAGCATATGCCAAATTTAAACGACACAGAAAGTCCCACTTTATGAATAC 600
 Db 660 AACACACGTTTCTAGCTTATGCCAGTTTAAAGACACAGGAGTGGCGCTCATTTGAATAC 719
 Qy 601 TCCTTCTACTCTGATGAGTCACTGCGATGCCAAGATCCAAAGACGTGACGGGTTCCATCCAAAG 660
 Db 720 TCCCTTCTATTCTGATGAGTCACTGCGATGCCAAGATCCCAAGACAGTGTGGATTCATACCCAAAG 779
 Qy 661 CGAGAGCTGTGAATCCAACTGTAAAGTCTTTTGTAAATACAGACTCTCTCAGCTCA 720
 Db 780 CGAGAGCTGTGAATCCAACTGTAAAGTCTTTTGTAAATATAGACTCTCTCAGCTCA 839
 Qy 721 GTCCAAATGCAACTTTCCATCAAAATCACTGCTCTCTGCTTCTATGTTGATAGGGGATCAC 780
 Db 840 TCCTCTAGTGGGCTCCCATCCAAATCCCTGCTCTGCTGCTGTCGCAATCTGTGGCAAGAGGGATCAC 899
 Qy 781 TACTTGTGTGATGAGACATGGGCAACAAAGAAAGATTTCTTGTGAGTGGCTCAGGAGG 840
 Db 900 TATTATGATGATGTTGGTGGGCTACAGAAAGAAAGATTTTCACTACAGTGGCTCAGGAGG 959
 Qy 841 ATTCAGACTATTTCGGTCAATGATTTTGTGACTATGATGATCCAGTGGAGATGGAC 900
 Db 960 ATTCAGACTATTTCGGTCAATGATTTTGTGACTATGATGATCAACCTTACGTTGGAAAC 1019
 Qy 901 TCGTTAGTGGCAGGCAACATTTGAAATGAGTACTTCTGCTGGGTTGGGAGATTTAGG 960
 Db 1020 TGTCCATCCGAGCAGCAGCATGTTGAAATGAGTACCAAGGCTGGGTGGGAGATTTAGG 1079
 Qy 961 CTTTCAAGAACCTTCAATTTTACCTTGTATGATTAATAGCTTCTCAAGATCATAGCAATGAA 1020
 Db 1080 CCCGCAAGAACCTTCAATTTCACTTCACTCTGATGGAAGCAGCTTCTATAAGATCATACGACAAA 1139
 Qy 1021 GAAGGTTACAGACATTTTGTCTTATTTCCAAATAGATTAAGAAAGCTGCAATTTATTACA 1080
 Db 1140 GATGGCTACAAACATCTGCCATTTCCGGAAGATAGAAAGACTGTACATTTATTACA 1199
 Qy 1081 AAAGGCACCTGGGAGTCTACGCGATGAGAGCTCTAACAGTGAATTTATCTATCTACATTT 1140
 Db 1200 AAAGGACCTTGGGAGTCTATGATATCGAAGCTCTGACCGAGATTTATCTATCTACTACATTT 1259

Qy 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAATCCAACTTATTGAC 1200
 Db 1260 AGTAACCAATATAAAGAAATGCCAGGAGGAAGAAATCTCTATAAATCAACTTACTGAC 1319
 Qy 1201 TATACAAAGTGACATGCCTCAGTTGTGAGCTGAATCCGAAAGGTGTCACTATTCT 1260
 Db 1320 CACACAAATGTGAAGTGCCTTTAGTTGTGACCTGAATCCAGAAAGATGTCACTATTATGCG 1379
 Qy 1261 GTGTCACTCAGTAAGAGGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCC 1320
 Db 1380 GTATCATTTAGTAAGAGGCAAGTACTATCAGCTGGGATGTTGGGGCCCCGGTCTGCC 1439
 Qy 1321 CTCTATATCTCTACACAGCAGCTGAATGATATAAAGGCTGAGAGTCTCTGGAAGCAATTC 1380
 Db 1440 CTCTACATCTCTACATCTGTAACCGATCATTAAGAGCTGCGAGTCTCTGGAAGCAATTTCT 1499
 Qy 1381 GCTTTGGATAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAACTGGACTTTCAATTT 1440
 Db 1500 GCTTTGGATAGATGCTGCAGGATGTCAGATGCCCTTCAAAAAAATTTGGACTTTCATTTGTT 1559
 Qy 1441 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATTAATCCAA 1500
 Db 1560 TTGAATGAAACAAAATTTTGGTATCAATGATCTTGGCCCCCTCATTTTGTATTAATCCAA 1619
 Qy 1501 AAATATCTCTCTACTATTAGATGTTGATGACGCCCATGTAGTCAAAAAAGCAGACACTGTC 1560
 Db 1620 AAATATCTCTCTACTATTAGATGTTATGACAGTCCCTGTAGTCAAAAAAGCAGACTGTTCC 1679
 Qy 1561 TTCAGACTGAACTGGGCCACTTACCTTGCAGACAGAAAAACATTTATATAGTACTAGCTTTT 1620
 Db 1680 TTCAGACTGAACTGGGCCACTTACCTTGCAGAGTACAGAAAAACATCATATAGTACTAGCTTT 1739
 Qy 1621 GATGCGAGAGAGTGGTTTACCAAGGAGATTAAGATCATGCAATCAACAGAGACTG 1680
 Db 1740 GACGCGACAGAGAGTGGTTTACCAAGGAGATTAAGATCATGCAATCAACAGAGAGTTG 1799
 Qy 1681 GGACATTTGAAGTTGAAGATCAAAATTCAGCAGCAGACAAATTTTCAAAAAATGGATTT 1740
 Db 1800 GGACATCTGGAAGTTGAAGATCAAAATTCAGCAGCAGGCAATTTTGTAAAAATGGGATTT 1859
 Qy 1741 GTGGACAAACAAACGAATTTGCGGCTGGTTCATATGAGAGGTACGTAACTCAATG 1800
 Db 1860 GTGGATACCAAGCGAGTTTGAATTTGGGCTGGTTCATATGAGGGGTATGTAACCTCAATG 1919
 Qy 1801 GTCTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCGTGGCGCTGTATCCCGG 1860
 Db 1920 GTCTCTGGGATCGGGAAGTGGGCTGTTCAAGTGTGCGGAATAGCTGTGGCACCTGTGTACGG 1979
 Qy 1861 TGGGAGTACTATGACTCAGTGTACACAAAGTTCATGAGGGTCTCCCAACTCCAGAGAC 1920
 Db 1980 TGGGAGTACTATGACTCAGTGTACACAGGGTTCATGAGGGTCTCCCAATTCAGAGAC 2039
 Qy 1921 AACCTTGACCAATTAAGAAATTTCAACAGTCAAGAGTGTGAAATTTTAAAAAAGTT 1980
 Db 2040 AACCTTGACCAATTAAGAAATTTCAACAGTCAAGAGTGTGAAATTTTAAAAAAGTT 2099
 Qy 1981 GAGTACCTCTCTTATTTATGGAACAGCAGATGATTAAGCTTTCACTTTTCAAGAGTCAAGT 2040
 Db 2100 GAGTACCTCTCTTATTTATGGAACAGCAGATGATTAAGCTTTCACTTTTCAAGAGTCAAGT 2159
 Qy 2041 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTTGTTATCTGATGAA 2100
 Db 2160 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTTGTTATCTGATGAA 2219
 Qy 2101 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160
 Db 2220 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2279
 Qy 2161 ATAAACCAATGTTTCTCTTTTAC 2182
 Db 2280 CTCCAGCAGCTGCTTCTCTTTAC 2301


```

RESULT 3
AY418368
LOCUS
DEFINITION Homo sapiens DPP4 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY418368
VERSION AY418368.1 GI:39774328
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 2193)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2193)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..2193
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>2193
/gene="DPP4"
/locus_tag="HCM6534"
ORIGIN
Query Match 70.9%; Score 1549.4; DB 10; Length 2193;
Best Local Similarity 77.3%; Pred. No. 0;
Matches 1550; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

Qy 80 ATCATGAATATCTCTACAAACAGAAATAATATCTTGGTATTCATCTGTAATCTGAATATGGAA 139
Db 188 ATCATGAATATCTCTACAAACAGAAATAATATCTTGGTATTCATCTGTAATCTGAATATGGAA 247
Qy 140 ACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTGGACATTCATCAATGATT 199
Db 248 ACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTGGACATTCATCAATGATT 307
Qy 200 ATTCATATCTCTGATGGGAGTTTATCTCTTAGAATACAACTACGTGAAGCAATGGA 259
Db 308 ATTCATATCTCTGATGGGAGTTTATCTCTTAGAATACAACTACGTGAAGCAATGGA 367
Qy 260 GGCATTCCTACACAGCTTCATATGACATTTATGATTTAAATAAAGGCGAGCTGATTACAG 319
Db 368 GGCATTCCTACACAGCTTCATATGACATTTATGATTTAAATAAAGGCGAGCTGATTACAG 427
Qy 320 AAGAGAGATTCCAAACACACAGTGGGTGCATATGATTTAAATAAAGGCGAGCTGATTACAG 379
Db 428 AAGAGAGATTCCAAACACACAGTGGGTGCATATGATTTAAATAAAGGCGAGCTGATTACAG 487
Qy 380 CATATGTTTGGAAACAAATGACATTTATGTTAAATAAAGGCGAGCTGATTACAGAA 439
Db 488 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 547
Qy 440 TCACATGACGGGAAAGAGATATATATATATGGAATAACTGACGTGGGTTTATGAG 499

```



```

/db_xref="taxon:9606"
/clone="CSOD1083YF07"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/clone="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match          47.7%; Score 1041.2; DB 1; Length 1081;
Best Local Similarity 98.7%; Pred. No. 3e-264;
Matches 1048; Conservative 10; Mismatches 3; Indels 1; Gaps 1

```

Query Match	47.7%	Score 1041.2;	DB 1;	Length 1081;
Best Local Similarity	98.7%	Pred. No. 3e-264;		
Matches 1048;	Conservative 10;	Mismatches 3;	Indels 1;	Gaps 1

Qy	1	AGTCGCAAAAC	TTTACAC	TCTAACTGATTTACTTAA	AAAAATAC	TATAGAC	TGGAAGTTATAC	60
Db	21	AGTCGCAAAAC	TTTACAC	TCTAACTGATTTACTTAA	AAAAATAC	TATAGAC	TGGAAGTTATAC	80
Qy	61	TCCTTTAAGAT	GGATTTTCAGAT	CATCAATATCTCTACA	ACAGAAAATAA	TATCTTGGTA	120	
Db	81	TCCTTTAAGAT	GGATTTTCAGAT	CATCAATATCTCTACA	ACAGAAAATAA	TATCTTGGTA	140	
Qy	121	TTCAATGCTGA	ATATGGAAA	CAGCTCAGTTTTCTTTGG	AACAGTACATTTGAT	GAGTTT	180	
Db	141	TTCAATGCTGA	ATATGGAAA	CAGCTCAGTTTTCTTTGG	AACAGTACATTTGAT	GAGTTT	200	
Qy	181	GGACATTCCTA	CTCAATGATTTATCA	ATATCTCTGATGGC	CAGTTTATCTCTTAGA	ATAC	240	
Db	201	GGACATTCCTA	CTCAATGATTTATCA	ATATCTCTGATGGC	CAGTTTATCTCTTAGA	ATAC	260	
Qy	241	AACCTACGTGA	AGCAATGAGGCATTC	CTACACAGCTTCATATGA	CATTTATGATTTAAAT	300		
Db	261	AACCTACGTGA	AGCAATGAGGCATTC	CTACACAGCTTCATATGA	CATTTATGATTTAAAT	320		
Qy	301	AAAAAGG	CAGCTGATTTACA	GAGAGAGAGATTC	CCAAAACA	CACACAGCTGGGT	360	
Db	321	AAAAAGG	CAGCTGATTTACA	GAGAGAGAGATTC	CCAAAACA	CACACAGCTGGGT	380	
Qy	361	CCAGTGGGT	CTATAAATTTGGC	CATATGTTTTGGAA	CAATATATATATATGGA	ATAT	420	
Db	381	CCAGTGGGT	CTATAAATTTGGC	CATATGTTTTGGAA	CAATATATATATATGGA	ATAT	440	
Qy	421	AAATTTACCA	AGTTACAGAAATCAC	ATGGA	CGGGAAAGAGATATATATATGGA	ATAT	480	
Db	441	AAATTTACCA	AGTTACAGAAATCAC	ATGGA	CGGGAAAGAGATATATATATGGA	ATAT	500	
Qy	481	ACTGACTGGG	TTATGAA	AGAGAGAGTCTTCA	GTCCTACTCTGCTCTGGTGGTCTCCA	540		
Db	501	ACTGACTGGG	TTATGAA	AGAGAGAGTCTTCA	GTCCTACTCTGCTCTGGTGGTCTCCA	560		
Qy	541	AAACGGCACT	TTTTTTAGCATATG	CCCCAAATTTAA	CGACACAGAGTCCC	CACTTATTCGAATAC	600	
Db	561	AAACGGCACT	TTTTTTAGCATATG	CCCCAAATTTAA	CGACACAGAGTCCC	CACTTATTCGAATAC	620	
Qy	601	TCCTTTCTACT	CTGATGAGTCACTG	CAGTACCCAAAGACTGTAC	CGGGTTCCATATATCCAAAG	660		
Db	621	TCCTTTCTACT	CTGATGAGTCACTG	CAGTACCCAAAGACTGTAC	CGGGTTCCATATATCCAAAG	680		
Qy	661	GCAGGAGCTGT	GAAATCCAACTGTA	AAAGTTCTTTGTTGTA	ATAATACAGACTCTCTCAGCTCA	720		
Db	681	GCAGGAGCTGT	GAAATCCAACTGTA	AAAGTTCTTTGTTGTA	ATAATACAGACTCTCTCAGCTCA	740		
Qy	721	GTCAACCAAT	GCACATTCATACAAAT	CAC	TGCTCTGCTTCTATGTTGATAGGGGATCAC	780		
Db	741	GTCAACCAAT	GCACATTCATACAAAT	CAC	TGCTCTGCTTCTATGTTGATAGGGGATCAC	800		
Qy	781	TACTTTGTGT	GATGTGACATGGG	CAACACAAGAAAGAAATTTCTTTG	CAGTGGCTCAGGAGG	840		
Db	801	TACTTTGTGT	GATGTGACATGGG	CAACACAAGAAAGAAATTTCTTTG	CAGTGGCTCAGGAGG	860		
Qy	841	ATTTCAGAA	CTATTTCGGT	CATGGATATTTGTG	CACTATGATGAATCCAGTGGG	AGATGGAAC	900	
Db	861	ATTTCAGAA	CTATTTCGGT	CATGGATATTTGTG	CACTATGATGAATCCAGTGGG	AGATGGAAC	920	

Qy	901	TGCTTAGTCGACGGCGAACATTTGAAATAGTACTCTGGCTGGGTTGGAAAGATTTAGG	960
Db	921	TGCTTAGTCGACGGCGAACATTTGAAATAGTACTCTGGCTGGGTTGGAAATTTWAGG	980
Qy	961	CCTTTCAGAAACCTCATTTTACCCCTTGATGGTAAATAGCTTCTACAAGATCATCAGCAATGAA	1020
Db	981	CCTTTCAGAAACCTCATTTTACCCCTTGATGGTAAATAGCTTCTACAAGATCATCAGCAATGAA	1040
Qy	1021	GAAGGTTACAGACACATTTTGCTATTTTCCAAATATAGATAAAAAA	1062
Db	1041	G-AGGTTTTCGACACMTTTTGCTATTTTCCAAATATAGATAAAAAA	1081

LOCUS	987 bp	mRNA	linear	EST 01-MAY-2004
BX399854	Homo sapiens PLACENTA	COT 25-NORMALIZED	Homo sapiens	cDNA
BX399854	clone CSDDI084YJ15	3-PRIME,	mRNA sequence.	

KEYWORDS EST. Homo sapiens (human)
SOURCE ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Hominidae; Homo.
1 (bases 1 to 987)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:306299915.

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five cloned
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
1082.1

FEATURES

source	1..987
Location/Qualifiers	

<http://www.genoscope.cns.fr/127/CGMAPS-CGVD1004/CGVDNF1GC-1092.1.1>

```

/db_xref="taxon:9606"
/clone="CS0D1084YJ15"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

Query Match 39.6%; Score 865; DB 5; Length 987;
Best Local Similarity 96.9%;
Pred. NO. 1.5e-217;
Matches 883; Conservative 8; Mismatches 19; Indels 1; Gaps 1;

1275	AGAGGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGCTCG - CCCCTCTATATCTCTAC	1333
Qy		
Db		
987	AGAGGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGCTCGCCCTCTATCTCTAC	928
Qy		
1334	ACAGCAGCTGAATGATAAAGGGCTGAGAGTCTCTGGAGACAATTCAGCTTTTGGATAAAA	1393
Qy		
Db		
927	ACAGCAGCTGAATGATAAAGGGCTGAGAGTCTCTGGAGACAATTCAGCTTTTGGATAAAA	868
Qy		
1394	TGCTGCAGAAATGTCAGATGCCCTCCAAAAAACTGGACTTCATTATTTTGAATGAACAA	1453
Qy		
Db		
867	TGCTGCAGAAATGTCAGATGCCCTCCAAAAAACTGGACTTCATTATTTTGAATGAACAA	808
Qy		
Db		

Qy	1920	CAACCTTGACCATTAACAGAAATTTCAACAGTCTATGAGCAGAGCTGAAATTTTAAACAAGT	1979
Db	660	CAACCTTTGACCATTTACAGAAATTTCAACAGTCTATGAGCAGAGCTGAAATTTTAAACAAGT	719
Qy	1980	TGAGTACCTCCTTTATTCATGGAACAGCAGAGATGATAGTTTCATTTTCAGCAGTCACTCA	2039
Db	720	TGAGTACCTCCTTTATTCATGGAACAGCAGAGATGATAGTTTCATTTTCAGCAGTCACTCA	779
Qy	2040	GATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGGATTTCCAGGCCAATGTGGTATACTGATGA	2099
Db	780	GATCTCCAGCCCTGGTTCGATGTTGGAGTGGGATTTCCAGGCCAATGTGGTATACTGATGA	839
Qy	2100	AGACCATGGAAATAGCTAGCAGCAGCAGCACACCAAC	2134
Db	840	AGACCATGG-ATAGCTAGCAGCAGCAGCACACCAAC	873

[illegible]

Qy	121	TTCAATGCTGAAATATGGAACACAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT	180
Db	231	TTCAATGCTGAAATATGGAACACAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT	290
Qy	181	GGACATCTATCAATGATTAATCAATATCTCCTGATGGGAGTTTATCTCTTAGAATAC	240
Db	291	GGACATCTATCAATGATTAATCAATATCTCCTGATGGGAGTTTATCTCTTAGAATAC	350
Qy	241	AACTACGTGAAGCAATGGAGGCATTCCTACACAGCTTCATATGACATTTATGATTTAAAT	300
Db	351	AACTACGTGAAGCAATGGAGGCATTCCTACACAGCTTCATATGACATTTATGATTTAAAT	410
Qy	301	AAAAGGCAGCTGATTAACAGAAGAGAGATTCCTCAACAAACACACAGTGGGTCAATGTCTCA	360
Db	411	AAAAGGCAGCTGATTAACAGAAGAGAGATTCCTCAACAAACACACAGTGGGTCAATGTCTCA	470
Qy	361	CCAGTGGGTTCATAAATGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGACCA	420
Db	471	CCAGTGGGTTCATAAATGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGACCA	530
Qy	421	AAATTTACCAAGTTACAGAAATCACATGGACGGGAAAAGAGATATAATATATATGGAATA	480
Db	531	AAATTTACCAAGTTACAGAAATCACATGGACGGGAAAAGAGATATAATATATGGAATA	590
Qy	481	ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA	540
Db	591	ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA	650
Qy	541	AACGGCACATTTTATGACATATGCCCCAATTTTAAACGACACAGAGTCCCACTTATGGAATAC	600
Db	651	AACGGCACATTTTATGACATATGCCCCAATTTTAAACGACACAGAGTCCCACTTATGGAATAC	710
Qy	601	TCCTCTTACTCTCATGAGTCACTGACAGTACCCAAAGACTGTACGGGTTCATATCCAAAG	660
Db	711	TCCTCTTACTCTCATGAGTCACTGACAGTACCCAAAGACTGTACGGGTTCATATCCAAAG	770
Qy	661	GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTTGTGTAATAACAGACTCTCTCAGCTCA	720
Db	771	GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTTGTGTAATAACAGACTCTCTCAGCTCA	830
Qy	721	GTCAACAATGCAACTTCCATACAAATCACTGCTCCTGCTTCTATGTTGATAGGGATCAC	780
Db	831	GTCAACAATGCAACTTCCATACAAATCACTGCTCCTGCTTCTATGTTGATAGGGAT-MC	889
Qy	781	TACTGTGTGATGTGACATGGGCAACACAAAGAAAGATTTCTTTGACGTGGCTCAGGAGG	840
Db	890	TACTGTGTGATGTGACATGGGCAACACAAAGAAAGATTTCTTTGACGTGGGCAAGGAGG	947
Qy	841	ATTCAGAACTATTTCGGTCAATGATATTTGACATATGATGAATCCAGTGGAGATGGAAC	900
Db	948	ATACAG-ACTATTTCGGTCAATGG-TAATTTGCMCTATGCTGTMT-CHGWRGGARATGGAM	1004
Qy	901	TGCTTAGTGG 910	
Db	1005	TRMTTAKTG 1014	
RESULT 10			
AL541949			
LOCUS			
DEFINITION			
AL541949 Homo sapiens PLACENTA Homo sapiens linear EST 24-MAR-2004			
5-PRIME, mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Homnidae; Homo.			
REFERENCE			
AUTHORS			
TITLE			
Full-length cDNA libraries and normalization			
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
1 (bases 1 to 1022)			

JOURNAL
COMMENT

Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:30546617.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1082.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DE007CG11Q1P1&c=1082.r.

FEATURES
source

1..1022
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE007YN21"
/tissue_type="PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 36.8%; Score 803.2; DB 1; Length 1022;
Best Local Similarity 94.5%; Pred. No. 3.5e-201;
Matches 843; Conservative 11; Mismatches 33; Indels 5; Gaps 2;

QY 1 AGTCGCAAACTTACACTCTAAGTATGTTTAAATAAATCTTATAGAGTGAAGTTATAC 60
DB 126 AGTCGCAAACTTACACTCTAAGTATGTTTAAATAAATCTTATAGAGTGAAGTTATAC 185
QY 61 TCTTAAAGATGATTTAGATCATGATATCTCTACAAACAGAAATATATCTTGGTA 120
DB 186 TCTTAAAGATGATTTAGATCATGATATCTCTACAAACAGAAATATATCTTGGTA 245
QY 121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAAACAGTACATTTGATGAGTTT 180
DB 246 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAAACAGTACATTTGATGAGTTT 305
QY 181 GGACATTTCTCAATGATTTTCAATATCTCTGATGGGCGAGTTTCTCTAGAAATAC 240
DB 306 GGACATTTCTCAATGATTTTCAATATCTCTGATGGGCGAGTTTCTCTAGAAATAC 365
QY 241 AACTAGCTGAAGCAATGGAGGCAATCTCTACAGCTTCATATGACATTTATGATTTAAAT 300
DB 366 AACTAGCTGAAGCAATGGAGGCAATCTCTACAGCTTCATATGACATTTATGATTTAAAT 425
QY 301 AAAAGGCGATGATTTACAGAGAGAGGATTTCCAAACACACAGTGGGTCACTGGTCA 360
DB 426 AAAAGGCGATGATTTACAGAGAGAGGATTTCCAAACACACAGTGGGTCACTGGTCA 485
QY 361 CCAAGTGGGTCAATAAATGGCATATGTTTGGAAACAAATGACATTTATGTTAAATTTGAACCA 420
DB 486 CCAAGTGGGTCAATAAATGGCATATGTTTGGAAACAAATGACATTTATGTTAAATTTGAACCA 545
QY 421 AATTTACCAAGTTTACAGATTCATGAGCGGGGAAAGAGATATATATATATGATGAATA 480
DB 546 AATTTACCAAGTTTACAGATTCATGAGCGGGGAAAGAGATATATATATATGATGHAT 605
QY 481 ACTGACTGGGTTTATCAAGAGGAGGCTTCAGTGCCTCTCTGCTGCTGCTGCTTCCCA 540
DB 606 ACTGACTGGGTTTATCAAGAGGAGGCTTCAGTGCCTCTCTGCTGCTGCTGCTTCCCA 665
QY 541 AACGGGCACTTTTATGATATGCCCCAATTTTAAACGACACAGAGTCCCACTTATTTGAATAC 600
DB 666 AACGGGCACTTTTATGATATGCCCCAATTTTAAACGACACAGAGTCCCACTTATTTGAATAC 725

QY 601 TCCCTTCTACTCTGATGAGTCACTGTCAGTACCCAAAGACTGTACGGGTTCCCATATCCAAAG 660
DB 726 TCCCTTCTACTCTGATGAGTCACTGTCAGTACCCAAAGACTGTACGGGTTCCCATATCCAAAG 785
QY 661 GCAGGAGCTGTGAATCCAACTGTAAGTCTTTTGTGTAATAACAGACTCTCTCTCAGCTCA 720
DB 786 GCAGGAGCTGTGAATCCAACTGTAAGTCTTTTGTGTAATAACAGACTCTCTCTCAGCTCA 845
QY 721 GTCACCAATGCAACTTCCATCAACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
DB 846 GTCACCAATGCAACTTCCATCAACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 905
QY 781 TACTTGTGTGATGTGACATGCGCAACACAAAGAAAGATTTCTTTCAGTGGTCTCAGGAGG 840
DB 906 TACTTGTGTGATGTGACATGCGG--CAMACAGAAATTTCTTGTGCTGGTCTCAGGAGG 962
QY 841 ATTACAGAACTATTCGGTCTGATGATATTTGTGATGATGAATCAGTGGAA 892
DB 963 --TTCAGACTATTCGGTCTGATATTTGTGATGATGATGATGATGATGATGATGATGATGATG 1012

RESULT 11
AL544456
LOCUS
DEFINITION
AL544456 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0I018YJ07 5-PRIME, mRNA sequence.
ACCESSION
VERSION
AL544456.3 GI:45744959
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 1007)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT
On Feb 15, 2001 this sequence version replaced gi:31266300.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
1082.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DI018CE04QP1&c=1082.r.

FEATURES
source

1..1007
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI018YJ07"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
1082.r

ORIGIN

Query Match 36.4%; Score 795.8; DB 1; Length 1007;
Best Local Similarity 95.4%; Pred. No. 3.2e-199;
Matches 852; Conservative 9; Mismatches 26; Indels 6; Gaps 4;

QY 1 AGTCGCAAACTTACACTCTAAGTATGTTTAAATAAATCTTATAGAGTGAAGTTATAC 60
DB 114 AGTCGCAAACTTACACTCTAAGTATGTTTAAATAAATCTTATAGAGTGAAGTTATAC 173


```

QY 61 TCCTTAAGATGGATTTTCAGATCATGATATCTCTCAAAACAGAAATAATATCTTGGTA 120
Db 174 TCCTTAAGATGGATTTTCAGATCATGATATCTCTCAAAACAGAAATAATATCTTGGTA 233
QY 121 TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATTTTGCAGTTT 180
Db 234 TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATTTTGCAGTTT 293
QY 181 GGACATTTCTATCAATGATTTATTCATATATCTCTGATGGCCAGTTTATTTCTTTAGATAC 240
Db 294 GGACATTTCTATCAATGATTTATTCATATATCTCTGATGGCCAGTTTATTTCTTTAGATAC 353
QY 241 AACTACGTGAAGCAATGAGGATTCCTACAGCTTCATATGACATTTATGATTTAAT 300
Db 354 AACTACGTGAAGCAATGAGGATTCCTACAGCTTCATATGACATTTATGATTTAAT 413
QY 301 AAAAGGCGAGCTGATTACAGAGAGAGAGGATTCCTCAAAACACACACAGTGGGTACATGCA 360
Db 414 AAAAGGCGAGCTGATTACAGAGAGAGGATTCCTCAAAACACACACAGTGGGTACATGCA 473
QY 361 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACATGACATTTATGTTAAATTTGAACCA 420
Db 474 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACATGACATTTATGTTAAATTTGAACCA 533
QY 421 AATTATCAAGTTACAGAAATCAGATGACGCGGGAAGAGATATATATATATATGAATAT 480
Db 534 AATTATCAAGTTACAGAAATCAGATGACGCGGGAAGAGATATATATATATATGAATAT 593
QY 481 ACTGACTGGGTTTATGAGAGAGAGTCTTCAAGTGCCTACTCTGCTCTGCTGGTCTTCCA 540
Db 594 ACTGACTGGGTTTATGAGAGAGAGTCTTCAAGTGCCTACTCTGCTCTGCTGGTCTTCCA 653
QY 541 AACGGCAGCTTTTATGACATATGCCCCAATTTAAACACACAGAGTCCCACTTTATTTGAATAC 600
Db 654 AACGGCAGCTTTTATGACATATGCCCCAATTTAAACACACAGAGTCCCACTTTATTTGAATAC 713
QY 601 TCCTTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 714 TCCTTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773
QY 661 GCAGGAGCTGGAATCCCAATGTAAGTTCTTTGTTGT -AAATACAGACTCTCTCAGCTC 719
Db 774 GCAGGAGCTGGAATCCCAATGTAAGTTCTTTGTTGT -AAATACAGACTCTCTCAGCTC 833
QY 720 AGTCACCAATGCACTTCCATACAAATCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 779
Db 834 AGTCACCAATGCACTTCCATACAAATCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 893
QY 780 CTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839
Db 894 CTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 950
QY 840 GATTCAGAGCTTTTGGTCAATGATATTTTGGATGATGATGATGATGATGATGATGATGATGAT 892
Db 951 G- -TTGAGATTTTGGTCTGATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1001

```

```

RESULT 12
LOCUS AUI22649 846 bp mRNA linear EST 01-AUG-2002
DEFINITION AUI22649 MAMMA1 Homo sapiens cdna clone MAMMA1002805 5', mRNA
sequence.
ACCESSION AUI22649
VERSION AUI22649.1 GI:10947365
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 846)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and

```

TITLE
JOURNAL
COMMENT

Isogai,T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES
source
1..846
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA1002805"
/tissue type="mammary gland"
/clone_lib="MAMMA1"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 35.2%; Score 768.4; DB 1; Length 846;
Best Local Similarity 98.4%; Pred. No. 5.6e-192;
Matches 797; Conservative 0; Mismatches 11; Indels 2; Gaps 2;
QY 1144 AATGAATATATAAGGAATGCCAGGAGGAAGGAATCTTTATATAAAATCCAACTTATTGACTAT 1203
Db 1 AATGAATATATAAGGAATGCCAGGAGGAAGGAATCTTTATATAAAATCCAACTTATTGACTAT 60
QY 1204 ACAAAGTGACATCCCTCAGTTGTGAGCTGAATCCGGAAGAGTGTCTAGTACTATTCTGTG 1263
Db 61 ACAAAGTGACATCCCTCAGTTGTGAGCTGAATCCGGAAGAGTGTCTAGTACTATTCTGTG 120
QY 1264 TCATTCTAGTAAAGGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCCCTC 1323
Db 121 TCATTCTAGTAAAGGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCCCTC 180
QY 1324 TATATCTTACACAGCAGCGTGAATGATAAGGGCTGAGAGTCTCTGGAGAGACAATTCAGCT 1383
Db 181 TATATCTTACACAGCAGCGTGAATGATAAGGGCTGAGAGTCTCTGGAGAGACAATTCAGCT 240
QY 1384 TTGGATAAATGCTGCAAGATGTCAGATGCTCCAGTCCCTCCAAATAAATCGAATTCATTTTGTG 1443
Db 241 TTGGATAAATGCTGCAAGATGTCAGATGCTCCAGTCCCTCCAAATAAATCGAATTCATTTTGTG 300
QY 1444 AATGAACAATAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGTATTAATCCAAAGAA 1503
Db 301 AATGAACAATAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGTATTAATCCAAAGAA 360
QY 1504 TATCTCTCTATTTAGATGTTATGAGGCGCCATGTAGTCAAAAAGCAGACACTGTCTTC 1563
Db 361 TATCTCTCTATTTAGATGTTATGAGGCGCCATGTAGTCAAAAAGCAGACACTGTCTTC 420
QY 1564 AGACTGAACTGGGCGCACTTACCTTGCAGGACACAGAAAACATTTATAGTAGTCTTGTAT 1623
Db 421 AGACTGAACTGGGCGCACTTACCTTGCAGGACACAGAAAACATTTATAGTAGTCTTGTAT 480
QY 1624 GGACAGAGAGTGGTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGAGACTGGGA 1683
Db 481 GGACAGAGAGTGGTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGAGACTGGGA 540
QY 1684 ACATTTGAACTGGAAGATCAAAATGGAAGCAGCAGACAATTTTCAAAAATGGGATTTGTG 1743
Db 541 ACATTTGAACTGGAAGATCAAAATGGAAGCAGCAGACAATTTTCAAAAATGGGATTTGTG 600
QY 1744 GACAACAAACGAATTTGCAATTTTGGGCTGCTATATGAGGGTACGTAACTCAATGGTC 1803
Db 601 GACAACAAACGAATTTGCAATTTTGGGCTGCTATATGAGGGTACGTAACTCAATGGTC 660

```

QY 1804 CTGGGATCGGGAAGTGGGTGTTCAAGTGTGGAATAGCGTGGCGCTGTATCCGGGTGG 1863
Db |||||
QY 661 CTGGGATCGGGAAGTGGGTGTTCAAGTGTGGAATACCGTGGCGCTGTATCCGGGTGG 720
Db |||||
QY 1864 GAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAGAAGACAAC 1923
Db |||||
QY 721 GAAGTCTATGACTCAATGT-CACAGAACGTTACATGGGTCTCCCAACTCCAGAAGACAAC 779
QY 1924 CTTGACCA-TTACAGAAATTCACAGTGCAT 1952
Db |||||
QY 780 CTTGACCAATTCGGGAAATTCACAGTGCAT 809
Db |||||

RESULT 13
BQ937942 900 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8922297 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6470282
DEFINITION 5', mRNA sequence.
ACCESSION BQ937942
VERSION BQ937942.1 GI:22353420
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14000 row: m column: 03
High quality sequence stop: 639.
Location/Qualifiers
1..900
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6470282"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2.1 Kb."

FEATURES
source
Query Match 33.1%; Score 722; DB 5; Length 900;
Best Local Similarity 98.4%; Pred. No. 1.le-179;
Matches 750; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 1294 CTGAGATGTTCCGGTCTCGTCCCTCTATCTACTACACAGCAGCGTGAATGATAA 1353
Db |||||
QY 89 CTGTGCTTTCAGTCTCGTCTCGTCCCTCTATCTACTACACAGCAGCGTGAATGATAA 148
Db |||||
QY 1354 GGGCTGAGAGTCTCGAAGACAATTCAGCTTTGGATATAAATGCTGCAGAAATGTCAGATG 1413
Db |||||
QY 149 GGGCTGAGAGTCTCGAAGACAATTCAGCTTTGGATATAAATGCTGCAGAAATGTCAGATG 208
QY 1414 CCTCCAAAAAATGAGCTTCATTATTTTGAATGAACAAAAATTTGGTATCAGATGATC 1473
Db |||||
QY 209 CCTCCAAAAAATGAGCTTCATTATTTTGAATGAACAAAAATTTGGTATCAGATGATC 268
QY 1474 TTGCTCTCTCATTTTGATAAAATCCAGAAATATCTCTACTATTAGATGTGTATGCAGGC 1533
Db |||||

```

```

Db 269 TTGCTCTCATTTTGTATAAAATCCAGAAATATCTCTACTATTAGATGTGTATGCAGGC 328
QY 1534 CCATGTAGTCAAAAAGCAGACACTGTCTTTCAGACTGAACCTGGCCACATTCCTTCGAAGC 1593
Db |||||
QY 329 CCATGTAGTCAAAAAGCAGACACTGTCTTTCAGACTGAACCTGGCCACATTCCTTCGAAGC 388
Db |||||
QY 1594 ACAGAAAACATTTATAGTAGCTAGCTTTTGGATGGCAGAGGAAGTGGTTTACCAAGGAGATAAG 1653
Db |||||
QY 389 ACAGAAAACATTTATAGTAGCTAGCTTTTGGATGGCAGAGGAAGTGGTTTACCAAGGAGATAAG 448
QY 1654 ATCATGTCATCAATCAACAGAGACTGGGAACATTTGAAGTTGAAGATCAAAATTCGAAGCA 1713
Db |||||
QY 449 ATCATGTCATCAATCAACAGAGACTGGGAACATTTGAAGTTGAAGATCAAAATTCGAAGCA 508
QY 1714 GCCAGACAATTTTCAAAAATGGGATTTTGTGACACAAACGAAATTTGGCGGTGG 1773
Db |||||
QY 509 GCCAGACAATTTTCAAAAATGGGATTTTGTGACACAAACGAAATTTGGCGGTGG 568
QY 1774 TCATATGGAGGCTACGTAACCTCAATGCTCTGGGATCGGGAAGTGGCGTGTTCGAAGTGT 1833
Db |||||
QY 569 TCATATGGAGGCTACGTAACCTCAATGCTCTGGGATCGGGAAGTGGCGTGTTCGAAGTGT 628
QY 1834 GGAATAGCCGTGGCGCTGTATCCCGTGGGAGTACTATGACTCAGTGTACACAGAACT 1893
Db |||||
QY 629 GGAATAGCCGTGGCGCTGTATCCCGTGGGAGTACTATGACTCAGTGTACACAGAACT 688
QY 1894 TACATGGGTCTCCCAACTCCAGAGACAACTTGACCACTTACAGAAATTCACAGCTCATG 1953
Db |||||
QY 689 TACATGGGTCTCCCAACTCCAGAGACAACTTGACCACTTACAGAAATTCACAGCTCATG 748
QY 1954 AGCAGAGCTGAAAAATTTTAAACAAAGTGTGAGTACCT-CCTTATTCATGGAACAGCAGATGA 2012
Db |||||
QY 749 AGCAGAGCTGAAAAATTTTAAACAAAGTGTGAGTACCTCCCTTATTCATGGAACAGCAGATGA 808
QY 2013 TAACTGTTTCACT-TTTCAGCAGTCACTCAGATCTCCAAAGCCC 2053
Db |||||
QY 809 TAACTGTTTCACTTTTCACTCAGTCACTCAGATCTCCAAAGCCC 850

RESULT 14
AL553051 897 bp mRNA linear EST 30-MAR-2004
LOCUS AL553051 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1072YM19 5-PRIME, mRNA sequence.
ACCESSION AL553051
VERSION AL553051.3 GI:45857821
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 897)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31274865.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the Not I and scor V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
1082.x
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0D1072AG10QPl&c=1082.r.
Location/Qualifiers
1..897
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source

```

```

/db_xref="taxon:9606"
/clone="CSODI072YM19"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 33.0%; Score 721; DB 1; Length 897;
Best Local Similarity 98.3%; Pred. No. 2.1e-179;
Matches 762; Conservative 7; Mismatches 2; Indels 4; Gaps 4;

Qy 1 AGTCGCAAACTTACACCTCACTAACTGATTACTTAAATACTTATAGACTGAAGTTATAC 60
Db |||
Qy 126 AGTCGCAAACTTACACCTCACTAACTGATTACTTAAATACTTATAGACTGAAGTTATAC 185
Db |||
Qy 61 TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAACAGAAATAATATCTTGGTA 120
Db |||
Qy 186 TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAACAGAAATAATATCTTGGTA 245
Db |||
Qy 121 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGACAGTACATTTGATGAGTTT 180
Db |||
Qy 246 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGACAGTACATTTGATGAGTTT 305
Db |||
Qy 181 GGACATCTTCAATGATTATCAATATCTCTGATGGGAGTTTATTTCTCTAGATATAC 240
Db |||
Qy 306 GGACATCTTCAATGATTATCAATATCTCTGATGGGAGTTTATTTCTCTAGATATAC 365
Db |||
Qy 241 AACTACGTGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTAAT 300
Db |||
Qy 366 AACTAGGTGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTAAT 425
Db |||
Qy 301 AAAAGGCAAGCTGATTACAGAGAGAGGATTCCTAAACACACACAGTGGGTACATGTC 360
Db |||
Qy 426 AAAAGGCAAGCTGATTACAGAGAGAGGATTCCTAAACACACACAGTGGGTACATGTC 485
Db |||
Qy 361 CCAGTGGTGCATAAATGGGCATATGTTTGGAAACATGACATTTATGTTAAATGAAAC 420
Db |||
Qy 486 CCAGTGGTGCATAAATGGGCATATGTTTGGAAACATGACATTTATGTTAAATGAAAC 545
Db |||
Qy 421 AATTACCAAGTTACAGAAATCACATGGAAGGGAAGGAGATATAATATAATGGAATA 480
Db |||
Qy 546 AATTACCAAGTTACAGAAATCACATGGAAGGGAAGGAGATATAATATAATGGAATA 605
Db |||
Qy 481 ACTGACTGGGTTTATGAAGAGGAGTCTTCACTGCTTACTCTGCTGCTGCTGCTGCT 540
Db |||
Qy 606 ACTGACTGGGTTTATGAAGAGGAGTCTTCACTGCTTACTCTGCTGCTGCTGCTGCT 665
Db |||
Qy 541 AACGGCACTTTTATGATATATGCCCAATTTAAACGACACAGAGTCCCACTTA-TTGA 599
Db |||
Qy 666 AACGGCACTTTTATGATATATGCCCAATTTAAACGACACAGAGTCCCACTTASTTGA 725
Db |||
Qy 600 CTCCTTCTACTCTGATGAGTCACTGCGAGTACCCAAAGACTGTACGGGTTCATATCCAA 659
Db |||
Qy 726 CTCCTTCTACTCTGATGAGTCACTGCGAGTACCCAAAGACTGTACGGGTTCATATCC 784
Db |||
Qy 660 GGCAGGAGCTGTGAATCCAACTGTAAGTTCTTTGTTGTAATAACAGACTCTCTCAGTCT 719
Db |||
Qy 785 GGCAGGAGCTGTGAATCCAACTGTAAGTTCTTTGTTGTTGT-AATACAGACTCTCTCAG 843
Db |||
Qy 720 AGTCACCAATGCAATCTTCCATACAAATCAGTCTCTGCTCTCTATGTTGATAGG 774
Db |||
Qy 844 AGTCACCAATGCAATCTTCCATACAAATCAGTCTCTCTATGTTGATAGG 897
Db |||

```

```

RESULT 15
LOCUS AUI139683 791 bp mRNA linear EST 05-AUG-2002
DEFINITION AUI139683 PLACE1 Homo sapiens cDNA clone PLACE1011107 5', mRNA
sequence.
ACCESSION AUI139683

```

```

AUI139683.1 GI:11001204
EST. Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 791)
AUTHORS Oca,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1..791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1011107"
/tissue type="placenta"
/clone_lib="PLACE1"
/notes="vector: pME18SFL3"

ORIGIN
Query Match 32.6%; Score 712.6; DB 1; Length 791;
Best Local Similarity 99.4%; Pred. No. 3.4e-177;
Matches 715; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1466 AGATGATCTTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1525
Db 8 AGATGATCTTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 67
Qy 1526 ATGCAGGCCCATGTAGTCAAAAAGCAGACACTGTCTTTCAGACTGAACTGGGCCACTTACC 1585
Db 68 ATGCAGGCCCATGTAGTCAAAAAGCAGACACTGTCTTTCAGACTGAACTGGGCCACTTACC 127
Qy 1586 TTGCAAGCAGACAAAAACATTTATAGTAGCTAGTTTGTATGGCAGAGGAGTGGTTACCAAG 1645
Db 128 TTGCAAGCAGACAAAAACATTTATAGTAGCTAGTTTGTATGGCAGAGGAGTGGTTACCAAG 187
Qy 1646 GAGTATAGATCATGTCATGCAATCAACAGAGACTGGGNAACATTTGAAAGTTGAAGTCAAA 1705
Db 188 GAGTATAGATCATGTCATGCAATCAACAGAGACTGGGNAACATTTGAAAGTTGAAGTCAAA 247
Qy 1706 TTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTTGTGGACAAACAAAGAAATTCGAAATTT 1765
Db 248 TTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTTGTGGACAAACAAAGAAATTCGAAATTT 307
Qy 1766 GGGGCTCGTCAATATGGAGGGTACGTAACCTCAATGGTCTCTGGGATCGGGAAGTGGCGGT 1825
Db 308 GGGGCTCGTCAATATGGAGGGTACGTAACCTCAATGGTCTCTGGGATCGGGAAGTGGCGGT 367
Qy 1826 TCAAGTGTGGAATACCGTGGCGCTGTATCCCGTGGGAGTACTATGACTCAGCTGTACA 1885
Db 368 TCAAGTGTGGAATACCGTGGCGCTGTATCCCGTGGGAGTACTATGACTCAGCTGTACA 427
Qy 1886 CAGAACGTTACATGGGTCTCCCACTCCAGAGCAACCTTGACCATTTACAGAAATTCAA 1945
Db 428 CAGAACGTTACATGGGTCTCCCACTCCAGAGCAACCTTGACCATTTACAGAAATTCAA 487
Qy 1946 CAGTCAATGAGCAGAGCTGAAATTTTAAACAAAGTTGAGTACCTCTCTTATTCATGAAACAG 2005
Db |||

```

Db 488 CAGTCATGAGCAGAGCTGAAATTTTAAACAAGTTGAGTACCTCCTTATTTCATGGAACAG 547
Qy 2006 CAGATGATAACGTTCACTTTTCAGCAGTCAGCTCAGATCTCCAAAGCCCTGGTCGATGTTG 2065
Db 548 CAGATGATAACGTTCACTTTTCAGCAGTCAGCTCAGATCTCCAAAGCCCTGGTCGATGTTG 607
Qy 2066 GAGTGGATTTCCAGGCAATGTGTATACCTGATGAAGACCATGGAATAGCTAGCAGCACAG 2125
Db 608 GAGTGGATTTCCAGGCAATGTGTATACCTGATGAAGACCATGGAATAGCTAGCAGCACAG 667
Qy 2126 CACACCAACATATATATACCCACATGAGCCACTTCATAAAACAATGTTTCTTTTACCT 2184
Db 668 CACACCAACATATATATACCCACATGAGCCACTTCATAAAACAATGTTTCTTTTACCTT 726

Search completed: February 17, 2006, 07:21:17
Job time : 8810 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2006, 05:06:42 ; Search time 654 Seconds
(without alignments)
7088.778 Million cell updates/sec

Title: US-10-659-055-2

Perfect score: 2184

Sequence: 1 agtcgcaaaacttacctct.....aacatgtttctttacct 2184

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7204252 seqs, 1061369211 residues

Total number of hits satisfying chosen parameters: 14408504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA New:*
- 1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
 - 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
 - 5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
 - 7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
 - 9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
 - 10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
 - 11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
 - 12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
 - 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2182.4	99.9	2301	8	US-10-522-789-1
2	1654.6	75.8	4852	12	US-11-136-527-2130
3	751	34.4	2814	9	US-11-245-147-168
4	751	34.4	2814	12	US-11-186-284-54
5	119.8	5.5	524	6	US-09-925-065A-71343
6	111.2	5.1	2819	12	US-11-136-527-2716
7	101.4	4.6	615	6	US-09-925-065A-899711
8	101.4	4.6	615	6	US-09-925-065A-866910
9	77.8	3.6	2649	12	US-11-151-601-21
10	77.8	3.6	3143	12	US-11-151-601-19
11	70.4	3.2	379	12	US-11-128-061-396
12	70.4	3.2	379	12	US-11-128-061-4038
13	70.4	3.2	379	12	US-11-128-049-396
14	70.4	3.2	379	12	US-11-128-049-4038
15	62.8	2.9	1464	6	US-09-925-065A-546409
16	54.4	2.5	634	6	US-09-925-065A-735709
17	43.2	2.0	125594	8	US-10-658-986-5
18	42.6	2.0	597	6	US-09-925-065A-106709
19	42	1.9	1790	6	US-09-925-065A-956
20	41.6	1.9	575	6	US-09-925-065A-858506

c	21	41.6	1.9	578	6	US-09-925-065A-857835	Sequence 857835,
c	22	40.8	1.9	577	6	US-09-925-065A-474460	Sequence 474460,
c	23	40.8	1.9	1568	8	US-10-750-185-62484	Sequence 62484, A
c	24	40.8	1.9	1568	8	US-10-750-623-62484	Sequence 62484, A
c	25	40.4	1.8	600	12	US-11-136-527-6812	Sequence 6812, Ap
c	26	40.4	1.8	1400	12	US-11-136-527-6364	Sequence 6364, Ap
c	27	40.4	1.8	2567	12	US-11-136-527-2268	Sequence 2268, Ap
c	28	39.8	1.8	628	6	US-09-925-065A-752885	Sequence 752885,
c	29	39.8	1.8	867	8	US-10-750-185-62362	Sequence 62362, A
c	30	39.8	1.8	867	8	US-10-750-185-62362	Sequence 62362, A
c	31	39.4	1.8	629	8	US-10-750-185-37439	Sequence 37439, A
c	32	39.4	1.8	629	8	US-10-750-623-37439	Sequence 37439, A
c	33	39	1.8	618	6	US-09-925-065A-884672	Sequence 884672,
c	34	39	1.8	152335	12	US-11-121-086-73	Sequence 73, Appl
c	35	38.8	1.8	574	6	US-09-925-065A-263454	Sequence 263454,
c	36	38.8	1.8	643	6	US-09-925-065A-950355	Sequence 950355,
c	37	38.8	1.8	1271	8	US-10-750-185-62485	Sequence 62485, A
c	38	38.8	1.8	1271	8	US-10-750-623-62485	Sequence 62485, A
c	39	38.4	1.8	574	6	US-09-925-065A-263451	Sequence 263451,
c	40	38.4	1.8	574	6	US-09-925-065A-263452	Sequence 263452,
c	41	38.4	1.8	574	6	US-09-925-065A-263453	Sequence 263453,
c	42	38.4	1.8	580	6	US-09-925-065A-404802	Sequence 404802,
c	43	38.4	1.8	610	6	US-09-925-065A-210331	Sequence 210331,
c	44	38.4	1.8	633	6	US-09-925-065A-323095	Sequence 323095,
c	45	38	1.7	101786	12	US-11-117-187-199	Sequence 199, App

ALIGNMENTS

RESULT 1

US-10-522-789-1
; Sequence 1, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2301)
; OTHER INFORMATION:
US-10-522-789-1

Query Match	99.9%	Score 2182.4;	DB 8;	Length 2301;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 2183;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	AGTCGCAAACTTACACTCTTACCTTAAATAATCTTATAGACTGAGTTATAC	60	
Db	115	AGTCGCAAACTTACACTCTTACCTTAAATAATCTTATAGACTGAGTTATAC	174	
Qy	61	TCCTTAGATGGATTTTCAGATCATGATATCTCTACAAACAAGAAATAATATCTTGTA	120	
Db	175	TCCTTAGATGGATTTTCAGATCATGATATCTCTACAAACAAGAAATAATATCTTGTA	234	
Qy	121	TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT	180	
Db	235	TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT	294	
Qy	181	GGACATTCTATCAATGATTTTCAATATCTCTGATGGCAGTTTATCTCTTGAATAC	240	
Db	295	GGACATTCTATCAATGATTTTCAATATCTCTGATGGCAGTTTATCTCTTGAATAC	354	

QY 241 AACTACGTGAAGCAATGGAGGCAATCCCTACACAGCTTCATATGACATATTTATGATTTAAAT 300
 Db |||||
 QY 355 AACTACGTGAAGCAATGGAGGCAATCCCTACACAGCTTCATATGACATATTTATGATTTAAAT 414
 Db |||||
 QY 301 AAAAGGCAGCTGATTAACAGAGAGAGGATTCAAAACAACACACAGTGGGTCAATGGTCA 360
 Db |||||
 QY 415 AAAAGGCAGCTGATTAACAGAGAGAGGATTCAAAACAACACACAGTGGGTCAATGGTCA 474
 Db |||||
 QY 361 CAGTGGGTCAATAATGGCATATGTTGGAAACATGACATTTATGTTAAAAATGAACCA 420
 Db |||||
 QY 475 CCAATGGGTCAATAATGGCATATGTTGGAAACATGACATTTATGTTAAAAATGAACCA 534
 Db |||||
 QY 421 AATTTACCAAGTTACAGAAATCACTGACGGGGAAGAGATATAATATAATGGAATA 480
 Db |||||
 QY 535 AATTTACCAAGTTACAGAAATCACTGACGGGGAAGAGATATAATATAATGGAATA 594
 Db |||||
 QY 481 ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGGCTACTCTGCTCTGFGTGGTCTCCA 540
 Db |||||
 QY 595 ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGGCTACTCTGCTCTGFGTGGTCTCCA 654
 Db |||||
 QY 541 AACGGCACTTTTATGACATATGCCAAATTTAAGACACACAGAGTCCCACTTATTTGAATAC 600
 Db |||||
 QY 655 AACGGCACTTTTATGACATATGCCAAATTTAAGACACACAGAGTCCCACTTATTTGAATAC 714
 Db |||||
 QY 601 TCCCTTCTACTCTGATGAGTCACTGACGATACCAAGAGCTGTACGGGTTCATATCCAAAG 660
 Db |||||
 QY 715 TCCCTTCTACTCTGATGAGTCACTGACGATACCAAGAGCTGTACGGGTTCATATCCAAAG 774
 Db |||||
 QY 661 GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTTGTAATAACACAGACTCTCTCAGCTCA 720
 Db |||||
 QY 775 GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTTGTAATAACACAGACTCTCTCAGCTCA 834
 Db |||||
 QY 721 GTACCAATGCAACTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGATCAC 780
 Db |||||
 QY 835 GTACCAATGCAACTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGATCAC 894
 Db |||||
 QY 781 TACTTGTGTGATGTGATGAGTGGCAACACAAAGAAAGAAATTTCTTGTGAGTGGCTCAGGAGG 840
 Db |||||
 QY 895 TACTTGTGTGATGTGATGAGTGGCAACACAAAGAAAGAAATTTCTTGTGAGTGGCTCAGGAGG 954
 Db |||||
 QY 841 ATTACAGAACTATTTCGGTCAATGATATTTGTGACTATGATGAATCCAGTGGGAAGATGGAAC 900
 Db |||||
 QY 955 ATTACAGAACTATTTCGGTCAATGATATTTGTGACTATGATGAATCCAGTGGGAAGATGGAAC 1014
 Db |||||
 QY 901 TGGTTAGTGGCAAGGCAACACATTTGAAATGAGTACTACTGGCTGGGTTCGAAAGATTAGG 960
 Db |||||
 QY 1015 TGGTTAGTGGCAAGGCAACACATTTGAAATGAGTACTACTGGCTGGGTTCGAAAGATTAGG 1074
 Db |||||
 QY 961 CCTTCAGAACCTCATTTTACCCTTGATGTTGATGATGATCTTCAAGATCATCAGCAATGAA 1020
 Db |||||
 QY 1075 CCTTCAGAACCTCATTTTACCCTTGATGTTGATGATGATCTTCAAGATCATCAGCAATGAA 1134
 Db |||||
 QY 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAGAGACTGCAATTTATTA 1080
 Db |||||
 QY 1135 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAGAGACTGCAATTTATTA 1194
 Db |||||
 QY 1081 AAGGCACTTGGGAAGTCACTGGGATAGAAAGTCTTAACAGATGATTTATCTATCTACTATT 1140
 Db |||||
 QY 1195 AAGGCACTTGGGAAGTCACTGGGATAGAAAGTCTTAACAGATGATTTATCTATCTACTATT 1254
 Db |||||
 QY 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTTATAAATCCAACTTATTTGAC 1200
 Db |||||
 QY 1255 AGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTTATAAATCCAACTTATTTGAC 1314
 Db |||||
 QY 1201 TATACAAAAGTGACATGCCCTCAGTGTGAGCTGAAATCCGGAAGAGTGTGAGTACTATTCT 1260
 Db |||||
 QY 1315 TATACAAAAGTGACATGCCCTCAGTGTGAGCTGAAATCCGGAAGAGTGTGAGTACTATTCT 1374
 Db |||||
 QY 1261 GTGTCAATTCAGTAAGAGGCAAGTATTTATCAGCTGAGATGTTCCGGTCTGGTCTGGCCC 1320
 Db |||||
 QY 1375 GTGTCAATTCAGTAAGAGGCAAGTATTTATCAGCTGAGATGTTCCGGTCTGGTCTGGCCC 1434
 Db |||||

QY 1321 CTCATATCTCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGACAATTCA 1380
 Db |||||
 QY 1435 CTCATATCTCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGACAATTCA 1494
 Db |||||
 QY 1381 GCTTTGGATAAATGCTCCAGAAATGTCAGATGCCCTCCAAAAGAACTGGACTTCAATTATT 1440
 Db |||||
 QY 1495 GCTTTGGATAAATGCTCCAGAAATGTCAGATGCCCTCCAAAAGAACTGGACTTCAATTATT 1554
 Db |||||
 QY 1441 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAAATCCAAAG 1500
 Db |||||
 QY 1555 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAAATCCAAAG 1614
 Db |||||
 QY 1501 AAATATCCTCTACTATTAAGATGTTGATGACAGGCCCATGTAGTCAAAAAGACAGACTGTC 1560
 Db |||||
 QY 1615 AAATATCCTCTACTATTAAGATGTTGATGACAGGCCCATGTAGTCAAAAAGACAGACTGTC 1674
 Db |||||
 QY 1561 TTGACACTGAATCTGGGCCACTTACCTTGCAGACACAGAAAACATTTATAGTAGTACTAGCTTT 1620
 Db |||||
 QY 1675 TTGACACTGAATCTGGGCCACTTACCTTGCAGACACAGAAAACATTTATAGTAGTACTAGCTTT 1734
 Db |||||
 QY 1621 GATGCGAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCGATGCAATCAACAGAGACTG 1680
 Db |||||
 QY 1735 GATGCGAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCGATGCAATCAACAGAGACTG 1794
 Db |||||
 QY 1681 GGAAATTTGAAAGTTGAAAGTCAAAATGAAAGCAGCAGACAATTTTCAAAAATGGGATTT 1740
 Db |||||
 QY 1795 GGAAATTTGAAAGTTGAAAGTCAAAATGAAAGCAGCAGACAATTTTCAAAAATGGGATTT 1854
 Db |||||
 QY 1741 GTGGACACAAACGAAATTTGGGGCTGGTCAATGAGGGGTGAGTAACTCAATG 1800
 Db |||||
 QY 1855 GTGGACACAAACGAAATTTGGGGCTGGTCAATGAGGGGTGAGTAACTCAATG 1914
 Db |||||
 QY 1801 GTCTCGGATCGGGAAGTGGGCTGTTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1860
 Db |||||
 QY 1915 GTCTCGGATCGGGAAGTGGGCTGTTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1974
 Db |||||
 QY 1861 TGGGAGTACTATGACTCAGTGTACACAAAGTGTACATGGGTCTCCCAACTCCAGAAAGAC 1920
 Db |||||
 QY 1975 TGGGAGTACTATGACTCAGTGTACACAAAGTGTACATGGGTCTCCCAACTCCAGAAAGAC 2034
 Db |||||
 QY 1921 AACCTTGACCAATTACAGAAATTTCAACAGTCAAGCAGCAGAGCTGAAATTTTAAACAAGTT 1980
 Db |||||
 QY 2035 AACCTTGACCAATTACAGAAATTTCAACAGTCAAGCAGCAGAGCTGAAATTTTAAACAAGTT 2094
 Db |||||
 QY 1981 GAGTACCTCTTATTTCATGGAACACAGATGATAAGTTCACCTTTGAGCAGTCAAGTCAAG 2040
 Db |||||
 QY 2095 GAGTACCTCTTATTTCATGGAACACAGATGATAAGTTCACCTTTGAGCAGTCAAGTCAAG 2154
 Db |||||
 QY 2041 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATCTGATGAA 2100
 Db |||||
 QY 2155 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATCTGATGAA 2214
 Db |||||
 QY 2101 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2160
 Db |||||
 QY 2215 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2274
 Db |||||
 QY 2161 ATAAACCAATGTTCTCTTTACCT 2184
 Db |||||
 QY 2275 ATAAACCAATGTTCTCTTTACCT 2298
 Db |||||

RESULT 2
 US-11-136-527-2130
 ; Sequence 2130, Application US/11136527
 ; Publication No. US20050287570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William M
 ; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
 ; FILE REFERENCE: 031896-041000 (AM101086)
 ; CURRENT APPLICATION NUMBER: US/11/136,527
 ; CURRENT FILING DATE: 2005-05-25
 ; PRIOR APPLICATION NUMBER: US 60/574,294

```

; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2.2
; SEQ ID NO: 2130
; LENGTH: 4852
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2130

```

Query Match	75.8%;	Score 1654.6;	DB 12;	Length 4852;
Best Local Similarity	84.7%;	Pred. No. 0;		
Matches 1850;	Conservative 16;	Mismatches 310;	Indels 9;	Gaps 1;
QY	7	AAAACCTTACACTCTTA	CTTA	CTTAA
DB	203	AGAACCTTACACACTA	KTCTGA	CTATTTAA
QY	67	AGATGGATTTTCAGAT	CAATGA	TATCTCTCA
DB	263	CGGTGGGGTTTCAGAT	TTCTGA	TACCTCTCA
QY	127	GCTGAATATGGAAAC	CAGCTC	AGTTTCTT
DB	323	GCTGAAACACGGGAA	CAGCTC	CAATTTCTT
QY	187	TCTATCAATGATATT	TCCAAT	TATCTCTG
DB	383	TCTATAGTGATTTAT	TTCAGT	GTCA
QY	247	GTGAAGCAATGGAGG	CAATTC	CTCACAG
DB	443	GTGAAGCAATGGAG	CACTCT	CA
QY	307	CAGCTGATACAGA	AGAGATTC	CA
DB	503	CAGCTGATCA	CAGAAG	AGAGATTC
QY	367	GGTCAATAAATTTG	CAATATG	TTTGG
DB	563	GGTCA	CA	MAATTTGG
QY	427	CCAAGTTACAGA	ATCA	ATCG
DB	623	CCTAGTCA	WAGGAY	CA
QY	487	TGGGTTATGA	AGAGAGTCTT	CA
DB	683	TGGGTTATGA	AGAGGAAATCTT	CG
QY	547	ACTTTTTTAC	ATAATGCC	CAATTTAA
DB	743	ACTTTTTCTAG	CTTATGCC	CAAGTTTAA
QY	607	TACTCTGAT	GAGTCAC	TG
DB	803	TACTCTGAT	GAGTCAC	TG
QY	667	GCTGTGAATCCA	ATGTGA	ATGTTCTT
DB	863	GCTGTGAATCCA	ATGTGA	ATGTTCTT
QY	727	AATGCAACTTC	CAATACAA	ATTCAG
DB	923	ACTACGATTC	CCCATG	CAATAC
QY	787	TGTGATGTGA	CTGAGG	CAAC
DB	983	TGTGACGTG	CCCTGGGTTT	CAGA
QY	847	AACATATTC	GGTCAT	TG
DB	1043	AACATATTC	GGTCAT	TG


```
Db 1696 TCGTTTGAATAATCCAGCTGCCTTAAGAGAAATTAAGAACTTGAAGTAGATGAAT 1755
Qy 1452 AAAATTTTGGTATCAGATGATCTTCCCTCCTCATTTTGATTAATCAAGAAATATCCCTCT 1511
Db 1756 TACTTTATGGTACAAGATGATCTTCCCTCCTCAATTTGACAGATCAAGAGATATCCCTT 1815
Qy 1512 ACTATTAGATGTATGAGGCCCAATGATGATCAAAAGAGACATGCTTTCAGACTGAA 1571
Db 1816 GCTAATTCAGGTATGTTGGTCCCTGCGAGTCAGAGTGTAAAGGTCTGTATTGCTGTTAA 1875
Qy 1572 CTGGSCCACTTACCTTGAAGCAGACAGAAACATATATAGTACGTAGCTTTGATGGCAGAG 1631
Db 1876 TTGGATATCTTATCTTCAAGTAAGAAAGGATGGTCAATGGCTTTGGTGGATGGTCGAGG 1935
Qy 1632 AAGTGGTTTCAAGGAGATAAGATCATGCGATGCAATCAACAGAGACTGGGAACATTTGA 1691
Db 1936 AACAGCTTTCCAGGTGACAACTCTCTATGCGAGTGATCGAAGACTGGGTGTTATGA 1995
Qy 1692 AGTTGAAGATCAAAATGGAAGCAGCAGACAAATTTTCAAAAATGGGATTTGGGACACAA 1751
Db 1996 AGTTGAAGACCAAGATTACAGCTGTGAGAAATTCATAGAAATGGGTTTCATTGATGAAA 2055
Qy 1752 ACRAATTCGAATTTGGGCTGGTCAATATGAGGGGTACGTAACTCAATGGTCTCTGGGATC 1811
Db 2056 AAGAAATAGCCATATGGGCTGGTCCCTATGAGGATACGTTTCATCACTGGCCCTTGCATC 2115
Qy 1812 GGGAGTGGCTGTTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGTGGGATGACTA 1871
Db 2116 TGGAACTGGTCTTTTCAAAATGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTA 2175
Qy 1872 TGACTCAGTGTACACAGAACGTTTACATGGGTCTCCCACTCCAGAGACAACTTTGACCA 1931
Db 2176 CGCGTCTGTACACAGAGAGATTCTGGGTCTCCCAACAAAGGATGATATCTTGAGCA 2235
Qy 1932 TTACAGAAATTCACAGTCTAGCAGCAGAGCTGAAATTTTAAACAGTTGAGTACCTCCT 1991
Db 2236 CTATAAGAAATTCACAGTGTGATGGCAAGAGCAGAAATATTTTCAGAAATGTAGACTATCTCT 2295
Qy 1992 TATTCATGGAACAGCAGATGATACGTTTCACTTTTCAGCAGTCACTGATCTCCAAAGC 2051
Db 2296 CATCCAGGAACAGCAGATGATATGTGCACTTTTCAAACTCAGCAGAGATGTCTAAAGC 2355
Qy 2052 CTTGGTCAATGTTGGAGTGGATTTCCAGGCAATGTTGATATCTGATGAAGACCACTGGAAT 2111
Db 2356 TCTGGTTAATGCACAGTGGATTTCCAGGCAATGTTGATCTCTGACCAAGNACCAGGCTT 2415
Qy 2112 AGCTAGCAGCAGCAGACACCAACATATATATATACCAATGAGCCACTTCATATAAACATG 2171
Db 2416 ATCCGGCTGTCCACGACAC---CACTTATACACCCACATGACCCACTTCTTAAAGCAGTG 2472
Qy 2172 TTTCTCTTT 2180
Db 2473 TTTCTCTTT 2481
```

RESULT 4
US-11-186-284-54
; Sequence 54, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM

```
; CURRENT APPLICATION NUMBER: US/11/186,284  
; CURRENT FILING DATE: 2005-07-21  
; PRIOR APPLICATION NUMBER: US/10/301,822  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 60/339,971  
; PRIOR FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: US 60/361,978  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/381,988  
; PRIOR FILING DATE: 2002-05-20  
; NUMBER OF SEQ ID NOS: 228  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 54  
; LENGTH: 2814  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (209)...(2491)  
US-11-186-284-54
```

```
Query Match 34.4%; Score 751; DB 12; Length 2814;  
Best Local Similarity 60.9%; Pred. No. 1.3e-170;  
Matches 1320; Conservative 0; Mismatches 825; Indels 24; Gaps 5;  
Qy 15 CACTCTAACTGATTACTTTAAAAAATACTTATAGACTGAAGTTATATCTCTTAAAGATGGAT 74  
Db 334 CACACTGAAGATATTTTAAATGGAACATTTCTTATATAAACATTTTTCACAACTGGAT 393  
Qy 75 TTCAGATCATGAATATCTCTACAAAGAAAAATAATATCTTGGTATTTCAATGCTGAATA 134  
Db 394 TTCAGGCAAGAAATATCTTCAATCAATCTGCAGATAACAATATAGTACTTTATAATATGA 453  
Qy 135 TGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTTGGACATTTCTATCAA 194  
Db 454 AACAGGACAATCATATACCATTTTGGATTAAGAACCAATGAAAGTGTGAATGCTTCAA - 512  
Qy 195 TGAATTTTCAATATCTCTGATGGCGAGTTTATTTCTCTTAGAATACAACTACGTGAAGCA 254  
Db 513 --ATTACGGCTTATCACCCTGATCGCAATTTGTATATCTAGAAAGTGATTTTCAAGCT 570  
Qy 255 ATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTAATAAAGAGCACTGAT 314  
Db 571 TTGAGATATCTCTTACACAGCAACATATATACATCTATGACCTTAGCAATGGAGAAATTTGT 630  
Qy 315 TACAGAGAGAGAGTATCCAAACACACACAGTGGGTACATGTCACAGTGGGTCTATAA 374  
Db 631 AAGAGGAAATGAGCTTCTCTGTCCAATTCAGTATTTATGCTGGTGGCTGTGGGAGTAA 690  
Qy 375 ATTGGCATATGTTTGGAAACAATGATCATTTATGTTAAAAATTTGAACCAAAATTTTACCAAGTTA 434  
Db 691 ATTAGCATATGCTATCAAAACATATCTATTGGAACAAAGACAGAGATCCACCTTT 750  
Qy 435 CAGAAATCAATGGACGGGGAAGAGATATAATATAATAGAAATACTGACTGGGTTTA 494  
Db 751 TCAAAATAACATTTAATGGAAGAAAAATAAATAATTTAATGGAATCCAGAGCTGGGTTTA 810  
Qy 495 TGAAGAGGAGTCTTTCAGTGCCTACTCTGCTCTGGTGGTCTCCAAAGCGCACTTTT 554  
Db 811 TGAAGAGGAAATGCTTCTTCAAAATATGCTCTCTGGTGGTCTCTTAAATGGAATAATTTT 870  
Qy 555 AGCATATGCCCAATTTTAAACGACACAGAGTCCCACTTATTGAATATCTCTTCTACTCTGA 614  
Db 871 GGCATATGGCGAATTTAATGATAGGATATACCAAGTATTGCTTATTTCTTATTATGGGA 930  
Qy 615 TGAGTCACTGCAATGCCAAAGACTGTACGGGTTTCCATATCCAAAGGAGAGAGCTGTGAA 674  
Db 931 TGA-----ACAAATATCTAGAACAAATAAATAATTTTAAATCCCAAGGCTGGAGCTAAGAA 984  
Qy 675 TCCAACTGTAAAGTCTTTGTTGTAATATACAGACTCTCTCAGCTCAGTCCCAATATGCAAC 734  
Db 985 TCCCGTTGTTGGGATATTTTATATCGATAC-----CACTTACCCCTGCGATGTAGG 1035
```

QY 735 TTCCATACAAATCACCTGCTCTTCTATGTATGTTAGGGGATCACTACTGTGTGATGT 794
Db 1036 TCCCAAGGAAGTGCCCTGTTCCAGCAATGATAGCCTCAAGTGAATATATTTTCAGTTGGCT 1095
QY 795 GACATGGGCAACAACAAGAAATTTCTTTGCAAGTGGCTCAGGAGGATTCAGAACTATTTC 854
Db 1096 CACGTGGGTACTGATGAACGAGTATGTTTGCAGTGGCTTAAAGAGAGTCCAGAAATGTTTC 1155
QY 855 GGTCAATGGATATTTGATGATATGATGAATCCAGTGGAGATGGAATGCTTTAGTGGCAGG 914
Db 1156 GGTCTCTGTCTATATGTGATTTTCAGGGAAGACTGGCAGACATGGGATTTGTCAAAGACCA 1215
QY 915 GCAACACATTTGAATAGATGACTTCTGCTGGGTGGAGATTTAGSCCTTCAGAACTTCA 974
Db 1216 GGAGCATATAGAAAGAACAGAACTGGATGGGCTGGTGGATTTCTTTGTTTCAAGACCACT 1275
QY 975 TTTTACCCCTTGATGTAATAGCTTTCTACAAGATCATCAGCAATAGAAAGGTTTACAGACA 1034
Db 1276 TTTGAGCTATGATGCCATTTCTGTAATACAAATATTTAGTACAAGGATGGCTTACAAACA 1335
QY 1035 CATTTGCTATTTCCAAATAGATAAAGAACTGCACATTTATTAACAAAGGACCTGGGA 1094
Db 1336 TATTCACATATATCAAGACACATGTGGAAATGCTATTTCAAAATTACAAAGTGGCAAGTGGGA 1395
QY 1095 AGTCATCGGGATAGAACTCTAACCAAGTATTTATCTATACATCATTTAGTAATGAATATAA 1154
Db 1396 GGCCATAAATATATTCAGAGTAACACAGGATTCACGTGTTTATTTAGCAATGAATTTGA 1455
QY 1155 AGGAATGCCAGGAGGAAGAAATCTTTATATAAATCCAACTTATTTGACTA---TACAAAAGT 1211
Db 1456 AGAATACCTCGAAGAAAGAAACATCTACAGAAATAGCAATGGAAAGCTATCCTCCAAGCAA 1515
QY 1212 GACATGCTCTCAGTTGTGACGTGAATCCGGAAGGTGTAGTACTATTCTGTGTCAATTCAG 1271
Db 1516 GAAGTGTGTACTTGGCATCTAAGGAAAGAAAGGTGCCAATATATACAGCAAGTTTTCAG 1575
QY 1272 TAAAGAGGGGAAGTATATCAGCTGAGATGTTCCGGTCTGTGCTGCCCTCTATACTCT 1331
Db 1576 CGACTACGCCAAGTACTATGCACTTGTCTGCTACGGCCAGGGATCCCAATTTCCACCT 1635
QY 1332 ACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGACAATTCAGCTTTGGATAA 1391
Db 1636 TCATGATGACGCACTGATCAAGAAATTAATCTCTGGAAGAAAACAAGAAATGGGAAA 1695
QY 1392 AATGCTGCAAGATGCTCAGATGCCCTCCTCAAAAACCTGGACTCATATTTTGAATGAAC 1451
Db 1696 TGCTTTGAAAAATATCCAGCTGCTTAAAGAGGAAATTAAGAAACCTTGAAGTAGATGAAT 1755
QY 1452 AAAATTTTGTGATCAGATGATCTTGCTCTCTCATTTTGTATAATCCAAGAAATATCCTCT 1511
Db 1756 TACTTTATGGTACAAGTGAATCTTCTCTCAATTTTGACAGATCAAGAAATATCCCTT 1815
QY 1512 ACTATTAGATGTGTATCAGGCCCCATGTAGTCAAAAAGCAGACACTGTCTTCAGACTGAA 1571
Db 1816 GCTAATTCAGTGTATGTTGTTGCTCTCTCTGCTGAGTGAAGTCTGTATTTGCTGTTAA 1875
QY 1572 CTGGGCCACTTACTCTTGAAGCAACAGAAATATATAGTAGCTAGCTTTGATGGCAGAGG 1631
Db 1876 TTGGATATCTTATCTTTCAGTAAGGAGGATGGTCAATTCCTTGGTGGATGGTGGAG 1935
QY 1632 AAGTGGTTACCAAGGAGATAAGATCATGATCAATCAACAGAACTGGCAACATTTTCA 1691
Db 1936 AACAGCTTTCCAGAGGTGACAAACTCTCTATGCAAGTGTATCGAAAGCTGGGTGTTATGA 1995
QY 1692 AGTTGAAGATCAAAATGAAGCAGCCAGACAATTTTCAAAAATGGGATTTTGTGACAAACA 1751
Db 1996 AGTTGAAGACCAAGATTACAGCTGTACAGAAATTTACAGAAATGGGTTTCATTGATGAAA 2055
QY 1752 ACGAATTTGCAATTTGGGGCTGGTCATATGAGGGTACGTAACTCAATGGTCTCGGATC 1811
Db 2056 AAGAATAGCCATATGGGGCTGGTCTTATGAGGATACGTTTTCATCATCGGCCCTTGCAATC 2115
QY 1812 GGGAGTGGCGTGTTCAGAGTGTGGAATAGCCGGGGCGCTGTATCCCGGTGGGAGTACTA 1871

Db 2116 TGGAACTGCTCTTTTCAAAATGGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTA 2175
QY 1872 TGACTCAGTGTACACAGAAAGCTTACATGGGTCTCCAACTCCAGAAAGACAACCTTGACCA 1931
Db 2176 CGCGTCTGTCTACACAGAGAGATTCATGGGTCTCCCAACAAAGGATGATTAATCTTTGAGCA 2235
QY 1932 TTACAGAAATCAACAGTCATGAGCAGAGCTGAAAATTTTAAACAAGTTGAGTACTCCT 1991
Db 2236 CTATAAGAAATTCAACTGTGATGGCAAGAGCAGAAATATTTTCAGAAATGTAGACTATCTTCT 2295
QY 1992 TATTTCATGGAACAGCAGATGATAACGTTTCACCTTTTCAGCAGTCAGCTCAGATCTCCAAAGC 2051
Db 2296 CATCCACGGAACAGCAGATGATAATGTGCACTTTTCAAAACTCAGCACAGATTTGCTTAAAGC 2355
QY 2052 CCGTGTCAATGTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAAGACATGGAAT 2111
Db 2356 TCTGGTTAATGCACAAGTGGATTTCCAGGCAATGTGGTACTCTGACCAAGAACCGGCTT 2415
QY 2112 AGCTAGCAGCAGCAGACACCAACATATATATACCCACATAGCCCACTTTCATAAAACAATG 2171
Db 2416 ATCCGGCTGTCCAGAAC---CACTTATACCCCACTATGACCCCACTTCTTAAGCAGTG 2472
QY 2172 TTTCTCTTTT 2180
Db 2473 TTTCTCTTT 2481

RESULT 5
US-09-925-065A-71343/c
; Sequence 71343, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71343
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-71343

Query Match 5.5%; Score 119.8; DB 6; Length 524;
Best Local Similarity 88.4%; Pred. No. 6.4e-19;
Matches 130; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 645 GGTTCATATCCAAAGGAGGAGCTGTAATCCAACTGTAAAAGTTCTTTGTTGTAATAC 704
Db 498 GATTTTCTTTCTAGGAGGAGCTGTGATCCAACTGTAAAAGTTCTTTGTTGTAATAC 439
QY 705 AGACTCTCTCAGCTCAGTCACCAATGCAACTTCCATACAAATCACTGCTCCTGCTTCTAT 764
Db 438 AGACTCTCTCAGCTCAGTCACCAATGCAACTTCCATACAAATCACTGCTCCTGCTTCTAT 379
QY 765 GTTGATAGGGATCACTACTTGTGTGA 791
Db 378 GTTGATAGGGTAAGACTCTTGTCTGA 352

RESULT 6

US-11-136-527-2716
; Sequence 2716, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2716
; LENGTH: 2819
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2716

Query Match 5.1%; Score 111.2; DB 12; Length 2819;

Best Local Similarity 46.6%; Pred. No. 1.6e-16; Mismatches 528; Indels 21; Gaps 3;
Matches 479; Conservative 0;

Qy	1160	TGCCAGGAGGAAGAACTTTATATAATCCAACTTATTGACTATACAAAAGTGACATGCC	1219
Db	1702	TGCCACGAAGACGACACCTCTACAGCGCCACACGGTAGACGCTTCAACAGGCAGTGTC	1761
Qy	1220	TCAGTTGTGAGCTGAATCCGGAAGGTGTGAGTACTATTCTGTGTCAATGTAAGAGG	1279
Db	1762	TGTCATGTGACCTG---GTGAGAACTGCACCTATGTCAAGTCTTCTTCACTATAACA	1818
Qy	1280	CGAAGTATTACAGCTGAGATGTTCCGCTCTGCTGCTGCCCTCTATCTTACACGCA	1339
Db	1819	TGGACTTCTTCTGCTCAAGTGTGAAGCCCTTGGCGTACCACCTGTTACTGTGACAATA	1878
Qy	1340	CGGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGCAAAATTCAGCTTTGGATAAAATGCTGC	1399
Db	1879	CCAGGATAGAGAAGATGTTTGACTTGGAGCAATGACGAGTGCAGAAACCATCT	1938
Qy	1400	AGAATGTCAGATCCCTCCAAAACCTGGACTTCATTTTGAATGAACAAATTTT	1459
Db	1939	ACGATCGCAGATGCTTAAATTCGAATACCGGAAATTTGAAGTGCAGATTTACAGCTTC	1998
Qy	1460	GGTATCAGATGATCTTGCTCTCTCATTTTGAATAATCCAGAAATATCTCTACTATTAG	1519
Db	1999	CAATGCAAACTTGAAGCCCGCCACCTTCACTGACACAGCCCACTACCCCTTTATTGCTCG	2058
Qy	1520	ATGTGTATGACGGCCCATGTAGTCMAAAGCAGACACTGTCTTCAGACTGAACCTGGGCA	1579
Db	2059	TGGTGGATGTACTTCAGGGAGTCAGAGCGTGTGAGAGGTTTGAAGTGAACCTGGGAGA	2118
Qy	1580	CTTACCTTGCAGACAGAAACATTTATAGTAGTACTTTTGTGCGAGAGGAAGTGTT	1639
Db	2119	CAGTGTGTTGTGAGCAGCATCGACAGTGTGGTCAAGTGTGACGGCGGAGCGGCT	2178
Qy	1640	ACGAGAGATAGATCATGATGCAATCAACAGAGACTGGGAACATTTGAAGTTGAAG	1699
Db	2179	TTACGGGAACCAAGCTCTTGACACGAAGTACGAGGCGGCTAGGCTTCTCTGGAGAGAAG	2238
Qy	1700	ATCAAAATTTGAAGCAGCAGACAAATTTTCAAAATTTGGGATTTGTGGACAAACAAATG	1759
Db	2239	ACCAGATGGAAGCTGTGGAACCAATGCTGAAGAGAGAGTACATTTGACAGACTCGGGTAG	2298
Qy	1760	CAATTTGGGGCTGCTCATATGAGGGTACGTAACCTCAATGGTCTCT-----GG	1807
Db	2299	CTGTGTTTGGGAAGGATTTATGTGGTACCTGAGTACTTACATCTCTCCAGCCAAAGGAG	2358
Qy	1808	GATCGGAGAGTGGCGTGTCAAGTGTGGAATAGCCGTGGCGCCCTGTATCCCGGTGGGAGT	1867
Db	2359	AAACCAAGGCCAGACTTTTACCTGCGGCTCTGCACTCTCTCCGATTAACGACTTCAAC	2418

Qy	1868	ACTATGACTCAGTGTACACAGACGGTTACATGGGTCTCCCAACTCCAGAGACACCTTG	1927
Db	2419	TCTATGCTCAGCATTTTCCGAGAGGTACCTTGGCTCCATGGACTCGCAACACAGCGT	2478
Qy	1928	ACCATACAGAAATTCACAGCTCATGACAGAGCTGAAAAATTTTAAACAAGTTGATACC	1987
Db	2479	ATGAGATGACCAAGCTGGCA-----CACAGAGTGTACAGACTGGAGGACCGAGCTTCC	2532
Qy	1988	TCCTTATTTCATGGAACAGCAGCATGATAACGTTTCACTTTTACAGCAGTCAGCTCAGATCTCCA	2047
Db	2533	TGATCATCCATGCTACGGCTGACGAAAAATTCATTTTCCAGCACACACAGAGCTCATCA	2592
Qy	2048	AAGCCCTGGTGCATGTTGGAGTGGATTTTCCAGGCAATGTGTATCTGATGAAGACCATG	2107
Db	2593	CGCAGCTAATCAAGGAAAGCCAAATTACAGCTTTACAGATATACCCAGATGAAGTCACT	2652
Qy	2108	GAATAGCTAGCAGCAGCAGCACCAACATATATATATATATATATATATATATATATATAT	2167
Db	2653	ACTTTACAGCGTGGCGCTCAAAACAGCACCTGTACAGGTCCATCATCGGATTTCTTTGTGG	2712
Qy	2168	AATGTTTC	2175
Db	2713	AATGTTTC	2720

RESULT 7

US-09-925-065A-899711/c
; Sequence 899711, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 899711
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-899711

Query Match 4.6%; Score 101.4; DB 6; Length 615;
Best Local Similarity 99.0%; Pred. No. 1.9e-14;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1771 TGGTCATATGAGGGTACGTAACCTCAATGGTCTCTGGATCGGAAGTGGCGTGTTCAG 1830
Db 478 TAGTCATATGAGGGTACGTAACCTCAATGGTCTCTGGATCGGAAGTGGCGTGTTCAG 419
Qy 1831 TGTGGATACCGTGGCGCTGTATCCCGTGGGAGTACTATG 1873
Db 418 TGTGGAATGCCGCGGCTGTATCCCGTGGGAGTACTATG 376

RESULT 8

US-09-925-065A-866910/c
; Sequence 866910, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 866910
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-866910

Query Match          4.6%; Score 101; DB 6; Length 615;
Best Local Similarity 98.1%; Pred. No. 2.3e-14; Indels 0; Gaps 0;
Matches 101; Conservative 1; Mismatches 1;

QY 1771 TGGTCATATGAGGGTACGTAACTCAATGTCCTGGGATCGGGAAGTGGCGTGTTCGAAG 1830
DB 478 TAGTCATATGAGGGTACGTAACTCAATGTCCTGGGATCGGGAAGTGGCGTGTTCGAAG 419

QY 1831 TGTGAATAGCCGTGGCGCTGTATCCCGTGGGAGTACTATG 1873
DB 418 TGTGAATAGCCGTGGCGCTGTATCCCGTGGGAGTATATG 376

RESULT 9
US-11-151-601-21
; Sequence 21, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE.
; FILE REFERENCE: MPI00-054PRCPIONDIVIM
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 866910
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-151-601-21

Query Match          3.6%; Score 77.8; DB 12; Length 2649;
Best Local Similarity 48.7%; Pred. No. 1.7e-08;
Matches 242; Conservative 0; Mismatches 252; Indels 3; Gaps 1;

QY 1619 TTGATGGCAGAGGAGTGGTTACCAAGGAGATAAGATCATGCATGCAATCAACAGAGAC 1678
DB 2057 TAGACACAGGGGATCCTGTACCAGGGCTTAAATTTGAAGGCCCTTTAAATATAAAA 2116
QY 1679 TGGGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGA---CAATTTTCAAAATGG 1735
DB 2117 TGGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAG 2176
QY 1736 GATTTGTGGACAAACAAAGAAATTGGGGCTGGTTCATATGAGGGGTACGTAACCT 1795
DB 2177 ATTTCAATTGACTTAGATCGTGTGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCC 2236
QY 1796 CAATGGTCTGGGATCGGGAAGTGGCGTGTTCAAAGTGGGAATAGCCGTGGCGCTGTAT 1855
DB 2237 TGATGGCAATTAATGACAGAGGTCAGATATCTTTCAGGGTTGCTATTTGGGGCCCAAGTCA 2296
QY 1856 CCGGTGGGAGTACTATGACTCAGTGTACACAGAAACGTTTACATGGGTCTTCCAACTCCAG 1915
DB 2297 CTCTGTGATCTTCTATGATACAGATACACGGAAGTTATATGGTTCACCTGACCAGA 2356
QY 1916 AAGCAACCTTGACCAATTCAGAAATTCACAGTCAAGCAGAGCTGAAAATTTTAAAC 1975
DB 2357 ATGAACAGGGCTATTACTTAGGATCTGTGGCCATCAAGACGAAAAGTTTCCCTCTGAAC 2416
QY 1976 AAGTTGAGTACTCTCTTATTCATGAAACAGCAGAGATGAACTTCACTTTTACAGCAGTCA 2035
DB 2417 CAATCGTTTACTGCTCTTACATGTTTCTCGATGGCTGGAAGCCATATGATTTACAGATCTATCCTC 2476
QY 2036 CTCAGATCTCCAAAGCCCTGGTCCGATGTTGGAGTGGATTTCCAGGCAATGTGTATCTG 2095
DB 2477 GTATATTACTGAGTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCCTC 2536
QY 2096 ATGAAGACCATGGAATA 2112
DB 2537 AGGAGAGACACAGCATA 2553

RESULT 10
US-11-151-601-19
; Sequence 19, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE.
; FILE REFERENCE: MPI00-054PRCPIONDIVIM
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
```

```
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(2874)
US-11-151-601-19

Query Match      3.6%; Score 77.8; DB 12; Length 3143;
Best Local Similarity 48.7%; Pred. No. 1.8e-08;
Matches 242; Conservative 0; Mismatches 252; Indels 3; Gaps 1;

Qy 1619 TTGATGCGAGGAGGAGTGGTTACCAAGGAGATAAGATCATGTCATCAACAGAGAC 1678
Db 2285 TAGACAACAGGGGATCTCTGTCCACGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAA 2344

Qy 1679 TGGGAACATTTGAAGTTGAAGATCAATTTGAAGCAGCAGA---CAATTTCAAAATGG 1735
Db 2345 TGGGTCAATAGAATTTGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATG 2404

Qy 1736 GATTTGTGGACAACAAGATTGCAATTTGGGCTGTGCATATGAGGGGTACGTAACTT 1795
Db 2405 ATTTCTAGTTAGATCGTGTGGCATCCACGGCTGTCTCTATGAGGATACCTCTCCC 2464

Qy 1796 CAATGTCCTGGGATCGGGAAGTGGCGTGTTCAGAGTGTGGAATAGCGGTGGCGCTGTAT 1855
Db 2465 TGATGGCATTAAATCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGCCCCAGTCA 2524

Qy 1856 CCGGTGGAGTACTATGACTCAGTGTACAGAAAGTTCATGGGTCTCCCAACTCCAG 1915
Db 2525 CTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTCAACCTGACCCAGA 2584

Qy 1916 AAGACAACCTTGACCATTTACAGAAATTCACAGTCATGAGCAGAGCTGAAAATTTTAAAC 1975
Db 2585 ATGAAAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAGTTCCCTCTGAAC 2644

Qy 1976 AAGTTGAGTACTCTCTTATTTCATGGAACAGCAGATGATAAGTTTCACTTTTCAGCAGTCAG 2035
Db 2645 CAATCGTTTACTGCTCTTACATGGTTCTCTGGATGAGATGTCCATTTTGCAATACCA 2704

Qy 2036 CTCAGATCTTCAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTG 2095
Db 2705 GTATATTACTGAGTTTGTAGTGAGGCTGGAAGGCCATATGATTTACAGATCTATCTCTC 2764

Qy 2096 ATGAGACCATGGAATA 2112
Db 2765 AGGAGACACAGCATATA 2781

RESULT 11
US-11-128-061-396
; Sequence 396, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4038
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-4038

Query Match      3.2%; Score 70.4; DB 12; Length 379;
Best Local Similarity 92.5%; Pred. No. 4.4e-07;
Matches 74; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1945 ACAGTCATGACGACGCTGAAATTTTAAACAAGTTGAGTACCTCTTATTCATGGAACA 2004
Db 2 ACAGTCATGACGACGCTGAAATTTTAAACAAGTTGAGTACCTCTTATTCATGGAACA 61

Qy 2005 GCAGATGATAACGTTCACTT 2024
Db 62 GCAGATGTTGAGTTTCAGTT 81

RESULT 12
US-11-128-061-4038
; Sequence 4038, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4038
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-4038

Query Match      3.2%; Score 70.4; DB 12; Length 379;
Best Local Similarity 92.5%; Pred. No. 4.4e-07;
Matches 74; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1945 ACAGTCATGACGACGCTGAAATTTTAAACAAGTTGAGTACCTCTTATTCATGGAACA 2004
Db 2 ACAGTCATGACGACGCTGAAATTTTAAACAAGTTGAGTACCTCTTATTCATGGAACA 61

Qy 2005 GCAGATGATAACGTTCACTT 2024
```

Db 62 GCAGATGCTGAGTTTCAGTT 81
||||| | | | | | | |

RESULT 13
US-11-128-049-396
; Sequence 396, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; PRIOR FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 396
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-049-396

Query Match 3.2%; Score 70.4; DB 12; Length 379;
Best Local Similarity 92.5%; Pred. No. 4.4e-07;
Matches 74; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1945 ACAGTCATGACGAGCTGAAAAATTTTAAACAAGTTGAGTACCTCTCTATTTCATGGAAACA 2004
Db 2 ACAGTCATGACGAGCTGAAAAATTTTAAACAAGTTGAGTACCTCTCTATTTCATGGAAACA 61
||||| | | | | | | |

QY 2005 GCAGATGATAACGTTCACTT 2024
Db 62 GCAGATGCTGAGTTTCAGTT 81
||||| | | | | | | |

RESULT 14
US-11-128-049-4038
; Sequence 4038, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; PRIOR FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4038
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-049-4038

Query Match 3.2%; Score 70.4; DB 12; Length 379;
Best Local Similarity 92.5%; Pred. No. 4.4e-07;
Matches 74; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1945 ACAGTCATGACGAGCTGAAAAATTTTAAACAAGTTGAGTACCTCTCTATTTCATGGAAACA 2004
Db 2 ACAGTCATGACGAGCTGAAAAATTTTAAACAAGTTGAGTACCTCTCTATTTCATGGAAACA 61
||||| | | | | | | |

QY 2005 GCAGATGATAACGTTCACTT 2024
Db 62 GCAGATGCTGAGTTTCAGTT 81
||||| | | | | | | |

RESULT 15
US-09-925-065A-546409/c
; Sequence 546409, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 546409
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-546409

Query Match 2.9%; Score 62.8; DB 6; Length 1464;
Best Local Similarity 69.7%; Pred. No. 5.3e-05;
Matches 85; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1752 ACGAATTCGAATTTGGGGCTGCTATATGAGGGTACGTAACCTCAATGGTCTGGGATC 1811
Db 756 ACTCATAACAATGTCTCTACAGTCTCTATGGAGGATACGTTTCATCACTGGCCCTTGCATC 697
||||| | | | | | | |

QY 1812 GGGAGTGCCTGTTTCAAGTGTGGAATAGCCCTGGCGCTGTATCCCGTGGGAGTACTA 1871
Db 696 TGGAACTGGTCTTTTCAATGTGATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTA 637
||||| | | | | | | |

QY 1872 TG 1873
Db 636 CG 635
|

Search completed: February 17, 2006, 06:44:30
Job time : 656 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2006, 05:01:13 ; Search time 1790 Seconds
(without alignments)
10089.565 Million cell updates/sec

Title: US-10-659-055-2

Perfect score: 2184

Sequence: 1 agtcgcaaaactacactct.....aacatgtttcttcttacct 2184

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134589005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2184	100.0	3407	5	US-10-002-593-5
2	2184	100.0	3407	5	US-10-165-603-6
3	2184	100.0	3407	6	US-10-423-714-5
4	2184	100.0	3407	6	US-10-295-027-921
5	2184	100.0	3407	7	US-10-794-899-40
6	2184	100.0	3407	10	US-11-041-674-5
7	2182.4	99.9	2301	6	US-10-295-027-589
8	2182.4	99.9	3445	9	US-10-952-459-17
9	2182.4	99.9	3445	9	US-10-956-157-741
10	2182.4	99.9	3445	9	US-10-631-467-42
11	2182.4	99.9	3913	9	US-10-887-553A-216
12	2180.8	99.9	2924	7	US-10-641-643-1026
13	1677.6	76.8	3316	9	US-10-631-467-969
14	1656.2	75.8	4835	3	US-09-917-800A-1570
15	1656.2	75.8	4835	5	US-10-165-603-5
16	1656.2	75.8	4835	7	US-10-794-899-39
17	752.2	34.5	2366	6	US-10-101-510-683
18	752.2	34.5	2788	6	US-10-269-909-22
19	752.6	34.5	2788	8	US-10-723-860-4170
20	752.6	34.5	2788	9	US-10-887-553A-235
21	752.6	34.5	2788	9	US-10-956-157-1251
22	752.6	34.5	3138	5	US-10-198-846-13171
23	751	34.4	2814	3	US-09-962-832-108

ALIGNMENTS

RESULT 1

; Sequence 5, Application US/10002593
; Publication No. US20020137120A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTING ENZYME INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Atty Docket No. US20020137120A1 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 3407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-002-593-5

Query Match	100.0%	Score 2184;	DB 5;	Length 3407;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2184;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGTCGCAAACTTACACTCTAACTGATTAACAAAATACCTTATAGACTGAAGTTATAC	60	
Db	190	AGTCGCAAACTTACACTCTAACTGATTAACAAAATACCTTATAGACTGAAGTTATAC	249	
Qy	61	TCCTTAGATGGATTTTCAGATCATGATATCTCTACAAACAGAAATATATCTTGGA	120	
Db	250	TCCTTAGATGGATTTTCAGATCATGATATCTCTACAAACAGAAATATATCTTGGA	309	
Qy	121	TTCAATCTCTGAATATGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGTAGAGTTT	180	
Db	310	TTCAATCTCTGAATATGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGTAGAGTTT	369	
Qy	181	GGACATTTCTATCAATGATTTATCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC	240	
Db	370	GGACATTTCTATCAATGATTTATCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC	429	
Qy	241	AACCTAGCTGAAGCAATGAGGCATTCCTACACAGCTTCATATGACATTTATGATTTAAAT	300	
Db	430	AACCTAGCTGAAGCAATGAGGCATTCCTACACAGCTTCATATGACATTTATGATTTAAAT	489	

QY 301 AAAAGCAGCTGATTACAGAGAGAGGATTCCAAACAAACACACAGTGGGTCAATGGTCA 360
DB |||||
QY 490 AAAAGCAGCTGATTACAGAGAGAGGATTCCAAACAAACACACAGTGGGTCAATGGTCA 549
DB |||||
QY 361 CCAAGTGGGTGATATAATGTCATATGTTGGAAACAATGACATTTATGTTAAAAATTGAACCA 420
DB |||||
QY 550 CCAAGTGGGTGATATAATGTCATATGTTGGAAACAATGACATTTATGTTAAAAATTGAACCA 609
DB |||||
QY 421 AATTTACCAAGTTACAGAAATCACTGGACGGGGAAGAAGATATATAATATTAATGGAATA 480
DB |||||
QY 610 AATTTACCAAGTTACAGAAATCACTGGACGGGGAAGAAGATATATAATATTAATGGAATA 669
DB |||||
QY 481 ACTGAGCTGGGTTATGAAGAGAGAGTCTTCAGTGGCTCTCTGCTCTGCTGCTGCTGCTCA 540
DB |||||
QY 670 ACTGAGCTGGGTTATGAAGAGAGAGTCTTCAGTGGCTCTCTGCTCTGCTGCTGCTGCTCA 729
DB |||||
QY 541 AACGGCAGCTTTTTTATGATATATGCCAAATTTAAACGACACAGAAAGTCCCACTTATTAATAT 600
DB |||||
QY 730 AACGGCAGCTTTTTTATGATATATGCCAAATTTAAACGACACAGAAAGTCCCACTTATTAATAT 789
DB |||||
QY 601 TCCCTTCTACTGATGAGTCACTGACGATACCCAAAGACTGTACGGGTTCCATATCCAAAG 660
DB |||||
QY 790 TCCCTTCTACTGATGAGTCACTGACGATACCCAAAGACTGTACGGGTTCCATATCCAAAG 849
DB |||||
QY 661 GCAGGAGCTGTAATCCAACTGTAAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA 720
DB |||||
QY 850 GCAGGAGCTGTAATCCAACTGTAAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA 909
DB |||||
QY 721 GTCCCAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTGATAGGGGATCAC 780
DB |||||
QY 910 GTCCCAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTGATAGGGGATCAC 969
DB |||||
QY 781 TACTTGTGTGATGATGAGTGGGCAACAAGAAAGAAATTTCTTTGCAAGTGGCTCAGGAG 840
DB |||||
QY 970 TACTTGTGTGATGATGAGTGGGCAACAAGAAAGAAATTTCTTTGCAAGTGGCTCAGGAG 1029
DB |||||
QY 841 ATTCCAGAACTATTCCGTGATGATATTTGTTGATGATGATGATGATGATGATGATGATGAT 900
DB |||||
QY 1030 ATTCCAGAACTATTCCGTGATGATATTTGTTGATGATGATGATGATGATGATGATGATGAT 1089
DB |||||
QY 901 TGCCTAGTGGCAGCGCAACACATTTGAAATGAGTACTACTGCTGGCTGGGTTGGAAAGATTAGG 960
DB |||||
QY 1090 TGCCTAGTGGCAGCGCAACACATTTGAAATGAGTACTACTGCTGGCTGGGTTGGAAAGATTAGG 1149
DB |||||
QY 961 CTTTCAGAACTCAATTTTACCCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB |||||
QY 1150 CTTTCAGAACTCAATTTTACCCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1209
DB |||||
QY 1021 GAAGGTTACAGACACATTTGCTATTTCCTAAATAGATAAAGAAAGACTGACATTTATTACA 1080
DB |||||
QY 1210 GAAGGTTACAGACACATTTGCTATTTCCTAAATAGATAAAGAAAGACTGACATTTATTACA 1269
DB |||||
QY 1081 AAAGGCACCTGGGAAGTCACTGGGATGAGAGCTCTAACAGATGATTTATCTATCTACATTT 1140
DB |||||
QY 1270 AAAGGCACCTGGGAAGTCACTGGGATGAGAGCTCTAACAGATGATTTATCTATCTACATTT 1329
DB |||||
QY 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTTATAAATCCAACTTATTGAC 1200
DB |||||
QY 1330 AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTTATAAATCCAACTTATTGAC 1389
DB |||||
QY 1201 TATACAAAGTGACATGCTCAGTCTGAGTGAATCCGGAAGGTTGAGTACTATTCT 1260
DB |||||
QY 1390 TATACAAAGTGACATGCTCAGTCTGAGTGAATCCGGAAGGTTGAGTACTATTCT 1449
DB |||||
QY 1261 GTGTCAATTCAGTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTGCTGCTGCC 1320
DB |||||
QY 1450 GTGTCAATTCAGTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTGCTGCTGCC 1509
DB |||||
QY 1321 CTCCTATCTACACAGCAGCGTGAATGATTAAGGCTGAGAGTCTCTGGAAGCAATTTCA 1380
DB |||||
QY 1510 CTCCTATCTACACAGCAGCGTGAATGATTAAGGCTGAGAGTCTCTGGAAGCAATTTCA 1569
DB |||||
QY 1381 GCTTTGGATAAATGCTGCAGAAATGTCCAGATGCCCTCCAAAAAATCGGACTTCAATTAT 1440
DB |||||

DB 1570 GCTTTGGATAAAAATGCTGCAGAAATGCTCCAGATGCCCTCAAAAAATCGACTTCATTATT 1629
QY |||||
DB 1441 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATTAATCCAG 1500
DB |||||
QY 1630 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATTAATCCAG 1689
DB |||||
QY 1501 AAATATCTCTACTATTAGATGTGTATGCAGGCCCATGTAGTCAAAAAAGCAGACACTGTC 1560
DB |||||
QY 1690 AAATATCTCTACTATTAGATGTGTATGCAGGCCCATGTAGTCAAAAAAGCAGACACTGTC 1749
DB |||||
QY 1561 TTCAAGCTGAACTGGGCCCACTTTACCTTGCAGACACAGAAAAATTTATAGTACTAGCTTT 1620
DB |||||
QY 1750 TTCAAGCTGAACTGGGCCCACTTTACCTTGCAGACACAGAAAAATTTATAGTACTAGCTTT 1809
DB |||||
QY 1621 GATGCGAGAGAAAGTGGTTTACCAAGAGATTAAGATCATGATGCAATCAACAGAGACTG 1680
DB |||||
QY 1810 GATGCGAGAGAAAGTGGTTTACCAAGAGATTAAGATCATGATGCAATCAACAGAGACTG 1869
DB |||||
QY 1681 GGAACATTTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAAATGGATTT 1740
DB |||||
QY 1870 GGAACATTTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAAATGGATTT 1929
DB |||||
QY 1741 GTGGACAAACAAACGAATTTGCAATTTTGGGGCTGGTCAATATGGAGGGTACGTAAACCTCAATG 1800
DB |||||
QY 1930 GTGGACAAACAAACGAATTTGCAATTTTGGGGCTGGTCAATATGGAGGGTACGTAAACCTCAATG 1989
DB |||||
QY 1801 GTCTGGGATCGGGAAGTGGGTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1860
DB |||||
QY 1990 GTCTGGGATCGGGAAGTGGGTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 2049
DB |||||
QY 1861 TGGGAGTACTATGACTCAGTGTACAGAACTTACATGCTGGGTCTCCCAACTCCAGAAAGAC 1920
DB |||||
QY 2050 TGGGAGTACTATGACTCAGTGTACAGAACTTACATGCTGGGTCTCCCAACTCCAGAAAGAC 2109
DB |||||
QY 1921 AACCTTGACCAATTTACAGAAATTTCAACAGTCTATGAGCAGAGCTGAAAAATTTTAAACAAGTT 1980
DB |||||
QY 2110 AACCTTGACCAATTTACAGAAATTTCAACAGTCTATGAGCAGAGCTGAAAAATTTTAAACAAGTT 2169
DB |||||
QY 1981 GAGTACTCTCTTATTCATGGAACAGCAGATGATTAAGTTTCACTTTTCCAGCAGTCAAGTCAAG 2040
DB |||||
QY 2170 GAGTACTCTCTTATTCATGGAACAGCAGATGATTAAGTTTCACTTTTCCAGCAGTCAAGTCAAG 2229
DB |||||
QY 2041 ATCTCCAAAGCCCTGCTGATGTTGGAGTGAATTTCCAGCAATGTGGTATATCTGATGAA 2100
DB |||||
QY 2230 ATCTCCAAAGCCCTGCTGATGTTGGAGTGAATTTCCAGCAATGTGGTATATCTGATGAA 2289
DB |||||
QY 2101 GACCATGGAATAGTACGACGACAGCAACCAACATATATATATATATATATATATATATATATAT 2160
DB |||||
QY 2290 GACCATGGAATAGTACGACGACAGCAACCAACATATATATATATATATATATATATATATATAT 2349
DB |||||
QY 2161 ATAAAAACAATGTTTCTCTTTTACCT 2184
DB |||||
QY 2350 ATAAAAACAATGTTTCTCTTTTACCT 2373
DB |||||

RESULT 2

US-10-165-603-6
; Sequence 6, Application US/10165603
; Publication NO. US20030021792A1
; GENERAL INFORMATION:
; APPLICANT: Roben, Paul W.
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
; FILE REFERENCE: TPTech.001A
; CURRENT APPLICATION NUMBER: US/10/165,603
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/297,021
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/305,117
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 33

QY 721 GTCCACCAATGCAACTTCCATACAAATCACTGCTCCTGCTTCTATGTTGATAGGGATCAC 780
DB 910 GTCCACCAATGCAACTTCCATACAAATCACTGCTCCTGCTTCTATGTTGATAGGGATCAC 969
QY 781 TACTTGTGTGATGTGACATGGGCAACACAAAGAAAGAAATTTCTTTGCAAGTGGCTCAGGAGG 840
DB 970 TACTTGTGTGATGTGACATGGGCAACACAAAGAAAGAAATTTCTTTGCAAGTGGCTCAGGAGG 1029
QY 841 ATTGAGAACTATTCGGTCAATGATATTTGTGATGATGATGAAATCCAGTGGAAAGATGGAAC 900
DB 1030 ATTGAGAACTATTCGGTCAATGATATTTGTGATGATGATGAAATCCAGTGGAAAGATGGAAC 1089
QY 901 TCTTTAGTGGCAGCGCAACACATTTGAAATGAGTACTTACCTGGCTGGTGGGATGATTTAGG 960
DB 1090 TGTCTAGTGGCAGCGCAACACATTTGAAATGAGTACTTACCTGGCTGGTGGGATGATTTAGG 1149
QY 961 CTTTCAGAACCTTCATTTTACCCTTGATGGTAATAGCTTCTCAAGATCATCAGCAATGAA 1020
DB 1150 CTTTCAGAACCTTCATTTTACCCTTGATGGTAATAGCTTCTCAAGATCATCAGCAATGAA 1209
QY 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCAATTTATTACA 1080
DB 1210 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCAATTTATTACA 1269
QY 1081 AAAGGACCTGGGAAGTCACTGGGATAGAGCTCTAACCAAGTATTTATCTATCTACTACATT 1140
DB 1270 AAAGGACCTGGGAAGTCACTGGGATAGAGCTCTAACCAAGTATTTATCTATCTACTACATT 1329
QY 1141 AGTAATGAATATAAGGAATGCCAGGAGGAAGAAATCTTTATAAAATCCAACTTATTGAC 1200
DB 1330 AGTAATGAATATAAGGAATGCCAGGAGGAAGAAATCTTTATAAAATCCAACTTATTGAC 1389
QY 1201 TATACAAAAGTGCATGCTCCTCAGTTGTGAGCTGAAATCCGGAAGGTGTGAGTACTATTCT 1260
DB 1390 TATACAAAAGTGCATGCTCCTCAGTTGTGAGCTGAAATCCGGAAGGTGTGAGTACTATTCT 1449
QY 1261 GTGCTATTCAGTAAAGAGCGGAATATATCAGCTGAGATGTTCCGGTCTCGCTGCGCC 1320
DB 1450 GTGCTATTCAGTAAAGAGCGGAATATATCAGCTGAGATGTTCCGGTCTCGCTGCGCC 1509
QY 1321 CTCTACTCTACACAGCAGCTGAATGATATAAGGCTGAGAGTCTCGGAAGCAATTCAC 1380
DB 1510 CTCTACTCTACACAGCAGCTGAATGATATAAGGCTGAGAGTCTCGGAAGCAATTCAC 1569
QY 1381 GCTTTGGATATAAATGCTGCAGAAATGTCAGATGCCCTCCAAAACCTGGAATTCATTAT 1440
DB 1570 GCTTTGGATATAAATGCTGCAGAAATGTCAGATGCCCTCCAAAACCTGGAATTCATTAT 1629
QY 1441 TTGAATGAACAAATAATTTTGGTATCAGATGATCTTGCCTCCTCATTTTGCATAAATCCAAG 1500
DB 1630 TTGAATGAACAAATAATTTTGGTATCAGATGATCTTGCCTCCTCATTTTGCATAAATCCAAG 1689
QY 1501 AAATATCCTCTACTATTAGATGTGATGAGGCCCCATGTAGTCAAAAAGCAGACACTGTC 1560
DB 1690 AAATATCCTCTACTATTAGATGTGATGAGGCCCCATGTAGTCAAAAAGCAGACACTGTC 1749
QY 1561 TTCAGACTGAATGGGCCACTTACCTTGCAGACACAGAAAAATTTATAGTAGTACTGTTT 1620
DB 1750 TTCAGACTGAATGGGCCACTTACCTTGCAGACACAGAAAAATTTATAGTAGTACTGTTT 1809
QY 1621 GATGGCAGGAAGTGGTTTACCAAGGAGATAGATCATGATGATGATGATGATGATGATGATG 1680
DB 1810 GATGGCAGGAAGTGGTTTACCAAGGAGATAGATCATGATGATGATGATGATGATGATGATG 1869
QY 1681 GGAACATTTTGAAGTGAAGATCAAAATTTGAAGCAGGCGAGACAAATTTTCAAAAATGGGATTT 1740
DB 1870 GGAACATTTTGAAGTGAAGATCAAAATTTGAAGCAGGCGAGACAAATTTTCAAAAATGGGATTT 1929
QY 1741 GTGGCAACAAACGAATTTGCAATTTTGGGGCTGGTCAATATGGAGGGTACGTAACTCAATG 1800
DB 1930 GTGGCAACAAACGAATTTGCAATTTTGGGGCTGGTCAATATGGAGGGTACGTAACTCAATG 1989

QY 1801 GTCTCGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCGTGGCGCTGTATCCCGG 1860
DB 1990 GTCTCGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCGTGGCGCTGTATCCCGG 2049
QY 1861 TGGGAGTACTATGACTCAGTGTACACAAAGTGTATCATGGGTCTCCCAACTCCAGAGAC 1920
DB 2050 TGGGAGTACTATGACTCAGTGTACACAAAGTGTATCATGGGTCTCCCAACTCCAGAGAC 2109
QY 1921 AACCTTGACCATACAGAAATTTCAACAGTGTATGAGCAGAGCTGGAATTTTAAACAAGTT 1980
DB 2110 AACCTTGACCATACAGAAATTTCAACAGTGTATGAGCAGAGCTGGAATTTTAAACAAGTT 2169
QY 1981 GAGTACCTCCTTATTTATGGAACAGCAGATGATTAACGTTTCACTTTTCAGCAGTCACTCAG 2040
DB 2170 GAGTACCTCCTTATTTATGGAACAGCAGATGATTAACGTTTCACTTTTCAGCAGTCACTCAG 2229
QY 2041 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2100
DB 2230 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2289
QY 2101 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160
DB 2290 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2349
QY 2161 ATAAAACAATGTTCTCTTTACCT 2184
DB 2350 ATAAAACAATGTTCTCTTTACCT 2373
RESULT 5
US-10-794-899-40
; Sequence 40, Application US/10794899
; Publication No. US20040146516A1
; GENERAL INFORMATION:
; APPLICANT: Utah Ventures
; FILE OF INVENTION: Lumen-Exposed Molecules and Methods Targeted Delivery
; FILE REFERENCE: 27110-715
; CURRENT APPLICATION NUMBER: US/10/794,899
; CURRENT FILING DATE: 2004-03-05
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 40
; LENGTH: 3407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-794-899-40
Query Match 100.0%; Score 2184; DB 7; Length 3407;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGTCGCAAACTTACACTTAACTGATTAATTAATAAATACTTATAGACTGAAGTTATAC 60
DB 190 AGTCGCAAACTTACACTTAACTGATTAATTAATAAATACTTATAGACTGAAGTTATAC 249
QY 61 TCCTTAAAGATGATTTTCAGATCATGAATATCTCTACAAACAAGAAAAATATATCTTGGTA 120
DB 250 TCCTTAAAGATGATTTTCAGATCATGAATATCTCTACAAACAAGAAAAATATATCTTGGTA 309
QY 121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTGATGAGTTT 180
DB 310 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTGATGAGTTT 369
QY 181 GGACATTTCTCAATGATTTTCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC 240
DB 370 GGACATTTCTCAATGATTTTCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC 429
QY 241 AACTACGTGAAGCAATGGAGGCAATTCCTACACAGCTTCTATATGACATTTTATGATTTAAAT 300
DB 430 AACTACGTGAAGCAATGGAGGCAATTCCTACACAGCTTCTATATGACATTTTATGATTTAAAT 489
QY 301 AAAAGCAGCTGATTAACAGAGAGGATTTCCAAACACACACAGTGGGTCACTGTC 360

490	AAAAGG	CAGCTGATTACAGAAAGAGAGGAGTTC	CCAAACAACA	CACAGTGGGT	CACATGTGCA	549
361	CCAGTGGG	CATAAATTTGGCATATGTTTGGAA	CAATGACATTTATGTTAAATTTGA	ACCA	420	
550	CCAGTGGG	CATAAATTTGGCATATGTTTGGAA	CAATGACATTTATGTTAAATTTGA	ACCA	609	
421	AAATTTACA	AGTTACAGAAATCACATGACGGG	GAAGAAGATATAATATATATATGGAATTA	480		
610	AAATTTACA	AGTTACAGAAATCACATGACGGG	GAAGAAGATATAATATATATGGAATTA	669		
481	ACTGAC	TGGGTTTATGAAGAGGAAGTCTT	CAGTGCCTTACTGCTCTCTGCTGGTGGTCTTCCA	540		
670	ACTGAC	TGGGTTTATGAAGAGGAAGTCTT	CAGTGCCTTACTGCTCTCTGCTGGTGGTCTTCCA	729		
541	AAOCCG	CACTTTTTTTAGCATATAGCCCAATTTAA	CCGACACAGAAGTCCCACTTTATTTGAATAC	600		
730	AAOCCG	CACTTTTTTTAGCATATAGCCCAATTTAA	CCGACACAGAAGTCCCACTTTATTTGAATAC	789		
601	TCCTTCT	ACTCTGATGAGTCACTGCACTGCCAAGACTGTA	CCGAAGA	CTGTACGGGTTCCATATFCCAAG	660	
790	TCCTTCT	ACTCTGATGAGTCACTGCACTGCCAAGACTGTA	CCGAAGA	CTGTACGGGTTCCATATFCCAAG	849	
661	GCAGG	AGCTGTGAATCCAACTGTAAAGTCTTTTGTGTA	ATAATACAGACTCTCTCAGTCTCA	720		
850	GCAGG	AGCTGTGAATCCAACTGTAAAGTCTTTTGTGTA	ATAATACAGACTCTCTCAGTCTCA	909		
721	GTCA	CAAAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTGTA	GATAGGGGATCAC	780		
910	GTCA	CAAAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTGTA	GATAGGGGATCAC	969		
781	TACTTGTG	TGATGTGACATGGGCAACAACAAGAAATTTCTTTG	CAGTGGCTCAGGAGG	840		
970	TACTTGTG	TGATGTGACATGGGCAACAACAAGAAATTTCTTTG	CAGTGGCTCAGGAGG	1029		
841	ATTTCAG	AACTTTTCGGTTCATGGATATTTTGTGACTATGATGA	TCCAGTGGAAAGTGAAC	900		
1030	ATTTCAG	AACTTTTCGGTTCATGGATATTTTGTGACTATGATGA	TCCAGTGGAAAGTGAAC	1089		
901	TGCTT	TAGTGCA	CGGCAACA	CATTTGAAATCAGTACTACTCGCTGGGTTGGAAGATTTAGG	960	
1090	TGCTT	TAGTGCA	CGGCAACA	CATTTGAAATCAGTACTACTCGCTGGGTTGGAAGATTTAGG	1149	
961	CCTT	CAGAA	CCTCATTTTACCCCTTGATGGTAAATAGCTTCTTACAAGATCATCAGCAATGAA	1020		
1150	CCTT	CAGAA	CCTCATTTTACCCCTTGATGGTAAATAGCTTCTTACAAGATCATCAGCAATGAA	1209		
1021	GAAG	TTACAGACATTTTGGTATTTTCCAAATAGATATAAAAGACTGCAATTTATPACA	1080			
1210	GAAG	TTACAGACATTTTGGTATTTTCCAAATAGATATAAAAGACTGCAATTTATPACA	1269			
1081	AAAGG	CACCTGGGAAGTTCATCGGATAGAAGCTCTACACAGTGCATATCTATCTACTACATTT	1140			
1270	AAAGG	CACCTGGGAAGTTCATCGGATAGAAGCTCTACACAGTGCATATCTATCTACTACATTT	1329			
1141	AGTAAT	TGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAATCCAACTTATTTGAC	1200			
1330	AGTAAT	TGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAATCCAACTTATTTGAC	1389			
1201	TATACAAA	AGTGACATGCCCTCAGTTGTGAGCTGAAATCCGGAAGAGGTGTGAGTACTATTTCT	1260			
1390	TATACAAA	AGTGACATGCCCTCAGTTGTGAGCTGAAATCCGGAAGAGGTGTGAGTACTATTTCT	1449			
1261	GTGT	CAATTCAGTAAAGAGGCGAAGTATATACAGTGCAGATGTTTCCGGTCTCGTCTGCGCC	1320			
1450	GTGT	CAATTCAGTAAAGAGGCGAAGTATATACAGTGCAGATGTTTCCGGTCTCGTCTGCGCC	1509			
1321	CTCT	TATCTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGGAAGACAATTTCA	1380			
1510	CTCT	TATCTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGGAAGACAATTTCA	1569			
1381	GCTTT	TGGATATAATGCTGCAAGATGTCCAGATGCCCTTCCAAAAATCTTGACTTCATTTATT	1440			
1570	GCTTT	TGGATATAATGCTGCAAGATGTCCAGATGCCCTTCCAAAAATCTTGACTTCATTTATT	1629			

RESIT.T 6

RESULT 6
IIS-11-041-674-5

US-11-041-674-3
: Sequence 5 Application IIS/11041674

; sequence 3; Application OS/II04
: Publication No. US20050181468A1

: PUBLICATION NO: US20

: APPLICANT: Vanderbilt University

APPLICANT: BROWN, Nancy J.

TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTING ENZYME

TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA

FILE REFERENCE: Atty Docket No. 1242/48/2/2/2

; CURRENT APPLICATION NUMBER: US/11/041,674

; CURRENT FILING DATE: 2005-01-24

;; PRIOR APPLICATION NUMBER: 60/244,524

;
; PRIOR FILING DATE: 2000-10-31

;
; PRIOR APPLICATION NUMBER: 10/002,593

;; PRIOR FILING DATE: 2001-10-31

; PRIOR APPLICATION NUMBER: 10/423,714

;
;
PRIOR FILING DATE: 2003-4-25

; NUMBER OF SEQ ID NOS: 10

QY 2041 ATCTCAAAGCCCTGGTGGATGTTGGAGTGGATTTCCAGGCAATGGGTATGATGAA 2100
 DB 2230 ATCTCAAAGCCCTGGTGGATGTTGGAGTGGATTTCCAGGCAATGGGTATGATGAA 2289
 QY 2101 GACCATTGAATAGCTAGCAGCAGCACACACCAACATATATATACCAATGAGCCACTTC 2160
 DB 2290 GACCATTGAATAGCTAGCAGCAGCACACACCAACATATATATATACCAATGAGCCACTTC 2349
 QY 2161 ATAAAACAATGTTCTCTTTTACCT 2184
 DB 2350 ATAAAACAATGTTCTCTTTTACCT 2373

RESULT 7

US-10-295-027-589
 ; Sequence 589, Application US/10295027
 ; Publication No. US20030232350A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevezi, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347,211
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/347,349
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 589
 ; LENGTH: 2301
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-295-027-589

Query Match 99.9%; Score 2182.4; DB 6; Length 2301;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AGTCGCAAACTTACACTCTAACTGATTATATAAATACTTATAGACTGAGTTATAC 60
 DB 115 AGTCGCAAACTTACACTCTAACTGATTATATAAATACTTATAGACTGAGTTATAC 174
 QY 61 TCCTTAGATGGATTGATGATGATATCTCTACAAACAGAAATAATATATCTGGTA 120

DB 175 TCCTTAGATGGATTGATGATGATATCTCTACAAACAGAAATAATATATCTGGTA 234
 QY 121 TTCAATGCTGAATATGGAACAGAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
 DB 235 TTCAATGCTGAATATGGAACAGAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT 294
 QY 181 GGACATTTCTATCAATGATTTTCAATATCTCTGATGGGAGTTTATTTCTTTAGAAATAC 240
 DB 295 GGACATTTCTATCAATGATTTTCAATATCTCTGATGGGAGTTTATTTCTTTAGAAATAC 354
 QY 241 AACTACGTGAAGCAATGGAGGCATTTCTTACACAGCTTCATATGACATTTATGATTTTAAAT 300
 DB 355 AACTACGTGAAGCAATGGAGGCATTTCTTACACAGCTTCATATGACATTTATGATTTTAAAT 414
 QY 301 AAAAGGCAGCTGATTTACAGAAAGAGAGGATTTCCAAACAAACACACAGTGGGTCCATGCTCA 360
 DB 415 AAAAGGCAGCTGATTTACAGAAAGAGAGGATTTCCAAACAAACACACAGTGGGTCCATGCTCA 474
 QY 361 CCAGTGGGTCAATAATTTGGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACCA 420
 DB 475 CCAGTGGGTCAATAATTTGGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACCA 534
 QY 421 AATTACCAAGTTACAGAAATCAGATGGAACGGGGAAGAGATATATATATATATGGAATA 480
 DB 535 AATTACCAAGTTACAGAAATCAGATGGAACGGGGAAGAGATATATATATATATGGAATA 594
 QY 481 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGCTGGTGGTCTCCA 540
 DB 595 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGCTGGTGGTCTCCA 654
 QY 541 AACGGCACTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 655 AACGGCACTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 714
 QY 601 TCCTTCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 715 TCCTTCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
 QY 661 GCAGGAGCTGTAATCCAACTGTAAGTTCTTCTGTTGTAATACAGACTCTCTCAGCTCA 720
 DB 775 GCAGGAGCTGTAATCCAACTGTAAGTTCTTCTGTTGTAATACAGACTCTCTCAGCTCA 834
 QY 721 GTCAACCAATGCAACTTCCATACAAACTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
 DB 835 GTCAACCAATGCAACTTCCATACAAACTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 894
 QY 781 TACTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 DB 895 TACTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954
 QY 841 ATTACAGAACTATTTCGGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 DB 955 ATTACAGAACTATTTCGGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1014
 QY 901 TGCTTAGTGGCAGGCAACCATTTGAAATGAGTACTACTGCTGGGTTGGAAGATTTAGG 960
 DB 1015 TGCTTAGTGGCAGGCAACCATTTGAAATGAGTACTACTGCTGGGTTGGAAGATTTAGG 1074
 QY 961 CCTTCAGAACTCTATTTTACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 DB 1075 CCTTCAGAACTCTATTTTACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1134
 QY 1021 GAGGTTTACAGACACATTTTCTTATTTCCAAATAGATAAAGAGTCCACATTTTATTTATACA 1080
 DB 1135 GAGGTTTACAGACACATTTTCTTATTTCCAAATAGATAAAGAGTCCACATTTTATTTATACA 1194
 QY 1081 AAAGGCCCTGGGAGGTCATCGGATAGAGGCTCTTAAACAGTATTATCTATCTACTACATT 1140
 DB 1195 AAAGGCCCTGGGAGGTCATCGGATAGAGGCTCTTAAACAGTATTATCTATCTACTACATT 1254
 QY 1141 AGTAATGATATATAAGGAATGCCAGGAGGAGGATCTTTTATAAAATCCAACTTATTTAGAC 1200
 DB 1255 AGTAATGATATATAAGGAATGCCAGGAGGAGGATCTTTTATAAAATCCAACTTATTTAGAC 1314

1201 TATACAAAAGTGACATGCTCAGTTGTGAGCTGAATCCGAAAGGTTGTCAGTACTATTCT 1260
Db TATACAAAAGTGACATGCTCAGTTGTGAGCTGAATCCGAAAGGTTGTCAGTACTATTCT 1374
Qy GTGTCAATTCAGTAAAGAGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCGTCTGCC 1320
Db GTGTCAATTCAGTAAAGAGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCGTCTGCC 1434
Qy CTCTATACCTACACAGCAGCGTGAATGATAAAGGCTGAGAGTCTCGAAGACATTTCA 1380
Db CTCTATACCTACACAGCAGCGTGAATGATAAAGGCTGAGAGTCTCGAAGACATTTCA 1494
Qy GCTTTGGATAAAATGCTGCAGAGTGTCCAGATGCCCTCCAAAACCTGACCTTCATTAT 1440
Db GCTTTGGATAAAATGCTGCAGAGTGTCCAGATGCCCTCCAAAACCTGACCTTCATTAT 1554
Qy TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGCCTCCTCATTTTGATAAATCCAAG 1500
Db TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGCCTCCTCATTTTGATAAATCCAAG 1614
Qy AAATATCCTTACTATTAGATGTGTATGAGGCCCATGTAGTCAAAAAGCAGACATGTC 1560
Db AAATATCCTTACTATTAGATGTGTATGAGGCCCATGTAGTCAAAAAGCAGACATGTC 1674
Qy TTCAGACTGAACCTGGGCCACTTACCTTGCAGACACAGAAAAACATTTATAGTAGTACTGTTT 1620
Db TTCAGACTGAACCTGGGCCACTTACCTTGCAGACACAGAAAAACATTTATAGTAGTACTGTTT 1734
Qy GATGCGAGAGGAGTGGTTTACCAAGGAGATAAGATCATGTCATGCAATCAACAGAGACTG 1680
Db GATGCGAGAGGAGTGGTTTACCAAGGAGATAAGATCATGTCATGCAATCAACAGAGACTG 1794
Qy GGAACTTTGAGTTGAAAGTCAAAATTTGAAGCAGCCAGACAATTTTCAAAAATGGGATTT 1740
Db GGAACTTTGAGTTGAAAGTCAAAATTTGAAGCAGCCAGACAATTTTCAAAAATGGGATTT 1854
Qy GTGGACAAACAAACGAATTCGAATTTGGGGCTGGTTCATATGGAGGGTACGTTAACTCAATG 1800
Db GTGGACAAACAAACGAATTCGAATTTGGGGCTGGTTCATATGGAGGGTACGTTAACTCAATG 1914
Qy GTCTGGGATCGGGAGTGGCGTGTTCAGTGTGGAAATAGCCGTGGCGCTGTATCCCGG 1860
Db GTCTGGGATCGGGAGTGGCGTGTTCAGTGTGGAAATAGCCGTGGCGCTGTATCCCGG 1974
Qy TGGGAGTACTATGACTCAGTGTACACAGAAAGTTCATATGGAGGGTACGTTAACTCAATG 1920
Db TGGGAGTACTATGACTCAGTGTACACAGAAAGTTCATATGGAGGGTACGTTAACTCAATG 2034
Qy AACCTTGACCATTTACAGAAATTCACAGTCATGAGCAGAGCTGAAAATTTTAAACAAGTT 1980
Db AACCTTGACCATTTACAGAAATTCACAGTCATGAGCAGAGCTGAAAATTTTAAACAAGTT 2094
Qy GAGTACTCTCTTATTCATGGAAACAGCAGATGATAAGTTTCACTTTTCCAGAGTCAAGTCAAG 2040
Db GAGTACTCTCTTATTCATGGAAACAGCAGATGATAAGTTTCACTTTTCCAGAGTCAAGTCAAG 2154
Qy ATCTCAAAGCCCTGTGATGTTGAGTGGATTTCCAGGCAATGGTGTATGATGAA 2100
Db ATCTCAAAGCCCTGTGATGTTGAGTGGATTTCCAGGCAATGGTGTATGATGAA 2214
Qy GACCATGGAATAGCTAGCAGCAGCAGCACCACATATATATATATATATATATATATATATAT 2160
Db GACCATGGAATAGCTAGCAGCAGCAGCACCACATATATATATATATATATATATATATATAT 2274
Qy ATAAAAAATGTTTCTCTTACCT 2184
Db ATAAAAAATGTTTCTCTTACCT 2298

RESULT 8

US-10-952-459-17

; Sequence 17, Application US/10952459

; Publication No. US20050074805A1
; GENERAL INFORMATION:
; APPLICANT: Kochan, Jarema Peter
; APPLICANT: Martin, Mitchell Lee
; APPLICANT: Robinski, James Andrew
; TITLE OF INVENTION: Specific Markers for Diabetes
; FILE REFERENCE: 21270US1
; CURRENT APPLICATION NUMBER: US/10/952,459
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 3445
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-952-459-17

Query Match 99.9%; Score 2182.4; DB 9; Length 3445;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGTCGCAAACTTACACTCTAACTGATTTACTTTAAAAAATCTATAGACTGAAGTATAC 60
Db 225 AGTCGCAAACTTACACTCTAACTGATTTACTTTAAAAAATCTATAGACTGAAGTATAC 284
Qy 61 TCCTTAAAGATGGATTTAGATCATCAATATCTCTACAAACAGAAAAATAATCTTGGTA 120
Db 285 TCCTTAAAGATGGATTTAGATCATCAATATCTCTACAAACAGAAAAATAATCTTGGTA 344
Qy 121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
Db 345 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 404
Qy 181 GGACATTTCTATCAATGATTTTCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC 240
Db 405 GGACATTTCTATCAATGATTTTCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC 464
Qy 241 AACTACGTGGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300
Db 465 AACTACGTGGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 524
Qy 301 AAAAGCAGCTGATTTACAGAGAGAGGATTCAAAACACACACAGTGGGTCACTGGTCA 360
Db 525 AAAAGCAGCTGATTTACAGAGAGAGGATTCAAAACACACACAGTGGGTCACTGGTCA 584
Qy 361 CCAGTGGGTTCATAAATTTGGCATAATGTTTGGACATGACATTTATGTTAAAAATCGAACCA 420
Db 585 CCAGTGGGTTCATAAATTTGGCATAATGTTTGGACATGACATTTATGTTAAAAATCGAACCA 644
Qy 421 AATTTTACCAGTTTACAGAAATCACATGGACGGGAAAGAGATATATATATATATATATAT 480
Db 645 AATTTTACCAGTTTACAGAAATCACATGGACGGGAAAGAGATATATATATATATATATAT 704
Qy 481 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGTGGGTCTCCA 540
Db 705 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGTGGGTCTCCA 764
Qy 541 AACGGCACTTTTTHAGCATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTATGGAATAC 600
Db 765 AACGGCACTTTTTHAGCATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTATGGAATAC 824
Qy 601 TCCTTCTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCATATCCAAAG 660
Db 825 TCCTTCTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCATATCCAAAG 884
Qy 661 GCAGGAGCTGTGAATCCCAACTGTAAAGTTCTTTTGTAAATACAGACTCTCTCAGCTCA 720
Db 885 GCAGGAGCTGTGAATCCCAACTGTAAAGTTCTTTTGTAAATACAGACTCTCTCAGCTCA 944
Qy 721 GTCACCAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTAGTAGGGATCAC 780
Db 945 GTCACCAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTAGTAGGGATCAC 1004

Db 585 ||||| CCAGTGGGTCATAAAATTGGCATATGTTTGGAAACAATGACATTTATGTTAAAATTGAACCA 644
QY 421 AATTTACCAAGTTTACAGAATTCACATGGACGGGAAAGAAAGATATAATATATAATGGAATA 480
Db 645 AATTTACCAAGTTTACAGAATTCACATGGACGGGAAAGAAAGATATAATATATAATGGAATA 704
QY 481 ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGCCCTACTCTGCTCTGTTGGTGGTCTCCA 540
Db 705 ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGCCCTACTCTGCTCTGTTGGTGGTCTCCA 764
QY 541 AACGGCAGCTTTTATAGCATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTTGAATAC 600
Db 765 AACGGCAGCTTTTATAGCATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTTGAATAC 824
QY 601 TCCTTCTACTGATGAGTCACTGACAGTACCCAAAGACTGTGACGGGTTCCATATCCAAAG 660
Db 825 TCCTTCTACTGATGAGTCACTGACAGTACCCAAAGACTGTGACGGGTTCCATATCCAAAG 884
QY 661 GCAGGAGCTGTAATCCAACTGTAAAGTCTTGTGTTGTAATAACAGACTCTCTCAGCTCA 720
Db 885 GCAGGAGCTGTAATCCAACTGTAAAGTCTTGTGTTGTAATAACAGACTCTCTCAGCTCA 944
QY 721 GTCACCAATGCCAATCTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGATCAC 780
Db 945 GTCACCAATGCCAATCTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGATCAC 1004
QY 781 TACTTGTGTGATGTGACATGGGCAACACAAAGAAAGAAATTTCTTTGACAGTGGCTCAGGAGG 840
Db 1005 TACTTGTGTGATGTGACATGGGCAACACAAAGAAAGAAATTTCTTTGACAGTGGCTCAGGAGG 1064
QY 841 ATTCAGAACTATTTCGGTCAATGATTTTGTGACTATGATGAATCCAGTGGAGATGGAC 900
Db 1065 ATTCAGAACTATTTCGGTCAATGATTTTGTGACTATGATGAATCCAGTGGAGATGGAAAC 1124
QY 901 TGTCTAGTGGCAGCGCAACACATTGAAATGAGTACTACTGGCTGGGTTGGAAAGATTAGG 960
Db 1125 TGTCTAGTGGCAGCGCAACACATTGAAATGAGTACTACTGGCTGGGTTGGAAAGATTAGG 1184
QY 961 CTTTCAGAACCTCATTTTACCCTTGATGTAATAGCTTCTCAAGATATCATGCAATGAA 1020
Db 1185 CTTTCAGAACCTCATTTTACCCTTGATGTAATAGCTTCTCAAGATATCATGCAATGAA 1244
QY 1021 GAAGTTACAGACACATTTGCTATTTCCAAATAGATAAAAGACTGACATTTATTACA 1080
Db 1245 GAAGSTTACAGACACATTTGCTATTTCCAAATAGATAAAAGACTGACATTTATTACA 1304
QY 1081 AAAGGCACCTGGGAAGTCACTCGGATAGAAAGCTCTAAACAGTGATTTATCTATACTACATT 1140
Db 1305 AAAGGCACCTGGGAAGTCACTCGGATAGAAAGCTCTAAACAGTGATTTATCTATACTACATT 1364
QY 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAATCCAACTTATTGAC 1200
Db 1365 AGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAATCCAACTTATTGAC 1424
QY 1201 TATACAAAAGTGNACATGCCCTCAGTTGTGAGCTGAATCCGGAAAGGTGTCAGTACTATTCT 1260
Db 1425 TATACAAAAGTGNACATGCCCTCAGTTGTGAGCTGAATCCGGAAAGGTGTCAGTACTATTCT 1484
QY 1261 GTGTCAATTCAATAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTGCTGCTGCC 1320
Db 1485 GTGTCAATTCAATAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTGCTGCTGCC 1544
QY 1321 CTCTATACTCTACACAGCAGCGTGAATAGATAAAGGGCTGAGAGTCTCTGGAAGCAATTCA 1380
Db 1545 CTCTATACTCTACACAGCAGCGTGAATAGATAAAGGGCTGAGAGTCTCTGGAAGCAATTCA 1604
QY 1381 GCTTTGGATATAATGCTGCAGATGTCCAGATGCCCTCCAAAAAATCGGACTTCATTATT 1440
Db 1605 GCTTTGGATATAATGCTGCAGATGTCCAGATGCCCTCCAAAAAATCGGACTTCATTATT 1664
QY 1441 TTGAATGAAACAAAAATTTTGGTATCAGATGATCTTGCCTCTCTATTTTGTATAAATCCAAG 1500

Db 1665 TTGAATGAAACAAAAATTTTGGTATCAGATGATCTTTGGCCTCTCTCATTTTGTATAAATCCAAG 1724
QY 1501 AAATATCCTCTACTATTAGATGTGTATGCAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1560
Db 1725 AAATATCCTCTACTATTAGATGTGTATGCAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1784
QY 1561 TTACAGCTGAACTCGGGCCACTTTACCTTGCAGGCACAGAAAAATTTATAGTAGTACTAGCTTT 1620
Db 1785 TTACAGCTGAACTCGGGCCACTTTACCTTGCAGGCACAGAAAAATTTATAGTAGTACTAGCTTT 1844
QY 1621 GATGCAGAGGAAGTGGTTTACCAAGGAGATGAATCATGATCATGCAATCAACAGAGACTG 1680
Db 1845 GATGCAGAGGAAGTGGTTTACCAAGGAGATGAATCATGATCATGCAATCAACAGAGACTG 1904
QY 1681 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1740
Db 1905 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGCAGACAAATTTTCAAAAATGGGATTT 1964
QY 1741 GTGGACAAACAACGAATTTGCAATTTGGGGCTGGTCAATATGAGAGGGTACGTAAACCTCAATG 1800
Db 1965 GTGGACAAACAACGAATTTGCAATTTGGGGCTGGTCAATATGAGAGGGTACGTAAACCTCAATG 2024
QY 1801 GTCTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCCGG 1860
Db 2025 GTCTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCCGG 2084
QY 1861 TGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAGAAGAC 1920
Db 2085 TGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAGAAGAC 2144
QY 1921 AACCTTGACCAATTACAGAAATTTCAACAGTCTATGACAGAGCTGAAAAATTTTAAAACAAGTT 1980
Db 2145 AACCTTGACCAATTACAGAAATTTCAACAGTCTATGACAGAGCTGAAAAATTTTAAAACAAGTT 2204
QY 1981 GAGTACCTCTTATTTCATGGAACAGCAGATGATAAAGTTTCACTTTTCAGCAGTCAAGCTCAG 2040
Db 2205 GAGTACCTCTTATTTCATGGAACAGCAGATGATAAAGTTTCACTTTTCAGCAGTCAAGCTCAG 2264
QY 2041 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTTCCAGGCAATGTTGTTATCTATGATGAA 2100
Db 2265 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTTCCAGGCAATGTTGTTATCTATGATGAA 2324
QY 2101 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2160
Db 2325 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2384
QY 2161 ATAAACCAATGTTTCTCTTTTACCT 2184
Db 2385 ATAAACCAATGTTTCTCTTTTACCT 2408

RESULT 10

US-10-631-467-42
; Sequence 42, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive
; FILE OF INVENTION: disease
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 3445
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-467-42

Db 2325 GACCATGGATAGCTAGCAGCAGCAGCACCAACATATATATATACCCACATGAGCCACTTC 2384
Qy 2161 ATAAACAATGTTCTCTTTACT 2184
Db 2385 ATAAACAATGTTCTCTTTACT 2408

RESULT 11
US-10-887-553A-216
; Sequence 216, Application US/10887553A
; Publication No. US20050085436A1
; GENERAL INFORMATION:
; APPLICANT: Garza, Dan
; APPLICANT: Li, Hao
; TITLE OF INVENTION: Method to treat conditions associated
; with insulin signalling dysregulation
; FILE REFERENCE: 4-33262
; CURRENT APPLICATION NUMBER: US/10/887,553A
; CURRENT FILING DATE: 2004-07-08
; PRIOR APPLICATION NUMBER: 60/485,883
; PRIOR FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 1208
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 216
; LENGTH: 3913
; TYPE: DNA
; ORGANISM: human
US-10-887-553A-216

Query Match 99.9%; Score 2182.4; DB 9; Length 3913;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AGTCGAAAACCTTACACTCTAACTGATTACTTAAATAAATACTTATAGACTGAAGTTATAC 60
Db 676 AGTCGAAAACCTTACACTCTAACTGATTACTTAAATAAATACTTATAGACTGAAGTTATAC 735
Qy 61 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAAACAAGAAAATATATCTTTGGTA 120
Db 736 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAAACAAGAAAATATATCTTTGGTA 795
Qy 121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGGAACAAGTACATTTTGCATGAGTTT 180
Db 796 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGGAACAAGTACATTTTGCATGAGTTT 855
Qy 181 GGACATTTCTATCAATGATTATCAATATCTCTGATGGGCAGTTTATCTCTTAGAATAC 240
Db 856 GGACATTTCTATCAATGATTATCAATATCTCTGATGGGCAGTTTATCTCTTAGAATAC 915
Qy 241 AACTAGCTGAAGCAATGGAGGCAATTCCTTACACAGCTTTCATATGACATTTATGATTTAAAT 300
Db 916 AACTAGCTGAAGCAATGGAGGCAATTCCTTACACAGCTTTCATATGACATTTATGATTTAAAT 975
Qy 301 AAAAGCAGCTGATTACAGAGAGGATTCCAAACAACACACAGTGGGTTCACATGGTCA 360
Db 976 AAAAGCAGCTGATTACAGAGAGGATTCCAAACAACACACAGTGGGTTCACATGGTCA 1035
Qy 361 CCAGTGGGTTCATAAAATGGGCATATGTTGGGAACAATGACATTTATGTTAAATTTGAACCA 420
Db 1036 CCAGTGGGTTCATAAAATGGGCATATGTTGGGAACAATGACATTTATGTTAAATTTGAACCA 1095
Qy 421 AATTTACCAAGTTTACAGATTCACATGGACGGGGAAGAAGATATATATATATGATTAATGAATA 480
Db 1096 AATTTACCAAGTTTACAGATTCACATGGACGGGGAAGAAGATATATATATATGATTAATGAATA 1155
Qy 481 ACTGACTGGGTTTATCAAGGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA 540
Db 1156 ACTGACTGGGTTTATCAAGGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA 1215
Qy 541 AACGGCACATTTTTTACCATATGCCCAATTTAACGACACAGAAAGTCCCACTTATTTGAATAC 600
Db 1216 AACGGCACATTTTTTACCATATGCCCAATTTAACGACACAGAAAGTCCCACTTATTTGAATAC 1275

Qy 601 TCCTTCTACTCTGATGAGTCACTGCAGTACCCAAAGATGTGTACGGTTCCCATATCCAAAG 660
Db 1276 TCCTTCTACTCTGATGAGTCACTGCAGTACCCAAAGATGTGTACGGTTCCCATATCCAAAG 1335
Qy 661 GCAGGAGCTGTGAATCCAACTGTAAGTCTTTTGTGTAAATACAGACTCTCTCAGCTCA 720
Db 1336 GCAGGAGCTGTGAATCCAACTGTAAGTCTTTTGTGTAAATACAGACTCTCTCAGCTCA 1395
Qy 721 GTCAACAAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTGATAGGGATCAC 780
Db 1396 GTCAACAAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTGATAGGGATCAC 1455
Qy 781 TACTTGTGTGATGTGACATGGGCAACAACAAGAAAGAAATTTCTTCGAGTGGCTCAGGAG 840
Db 1456 TACTTGTGTGATGTGACATGGGCAACAACAAGAAAGAAATTTCTTCGAGTGGCTCAGGAG 1515
Qy 841 ATTCAAGAACTATTTCGGTCAATGATATTTGTGACTATGATGAATCCAGTGGAAAGATGGAAC 900
Db 1516 ATTCAAGAACTATTTCGGTCAATGATATTTGTGACTATGATGAATCCAGTGGAAAGATGGAAC 1575
Qy 901 TGCTTAGTGGCACGGCAACACATTCGAAATGAGTACTACTGGCTGGGTTGGAAAGATTTAGG 960
Db 1576 TGCTTAGTGGCACGGCAACACATTCGAAATGAGTACTACTGGCTGGGTTGGAAAGATTTAGG 1635
Qy 961 CCTTCAGAACCTCATTTTACCCTTGATGGTAAATAGCTTCTCAAGATCATCAGCAATGAA 1020
Db 1636 CCTTCAGAACCTCATTTTACCCTTGATGGTAAATAGCTTCTCAAGATCATCAGCAATGAA 1695
Qy 1021 GAAGGTTACAGACACATTTTGCTATTTCAAAATAGATAAAAAGACTGCACATTTATTACA 1080
Db 1696 GAAGGTTACAGACACATTTTGCTATTTCAAAATAGATAAAAAGACTGCACATTTATTACA 1755
Qy 1081 AAAGCACTCGGAAAGTCAATCGGATAGAGCTCTAACAGTGAATATCTATACATACATT 1140
Db 1756 AAAGCACTCGGAAAGTCAATCGGATAGAGCTCTAACAGTGAATATCTATACATACATT 1815
Qy 1141 AGTAATGAATATAAGGAATGCCAGGAGGAAGGAATCTTTATAAATCCAACTTTATTGAC 1200
Db 1816 AGTAATGAATATAAGGAATGCCAGGAGGAAGGAATCTTTATAAATCCAACTTTATTGAC 1875
Qy 1201 TATACAAAAGTGACATGCTCAGTTGTGAGCTGAATCCGGAAAGGTGTGAGTACTATTCT 1260
Db 1876 TATACAAAAGTGACATGCTCAGTTGTGAGCTGAATCCGGAAAGGTGTGAGTACTATTCT 1935
Qy 1261 GTGTCATTTCAGTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCGTCTGCC 1320
Db 1936 GTGTCATTTCAGTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCGTCTGCC 1995
Qy 1321 CTCATATCTTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGACAATTCA 1380
Db 1996 CTCATATCTTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGACAATTCA 2055
Qy 1381 GCTTTGGATATAAATCTCTCAGAAATGTCAGATGCTCCCTCCAAAAAACTGGACTTTATT 1440
Db 2056 GCTTTGGATATAAATCTCTCAGAAATGTCAGATGCTCCCTCCAAAAAACTGGACTTTATT 2115
Qy 1441 TTGAATAGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCTATTTGATATAAATCCAAG 1500
Db 2116 TTGAATAGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCTATTTGATATAAATCCAAG 2175
Qy 1501 AAATATCTCTTACTATTAGATGTTATGCGGGCCCATGTAGTCAAAAAGCAGACACTGTC 1560
Db 2176 AAATATCTCTTACTATTAGATGTTATGCGGGCCCATGTAGTCAAAAAGCAGACACTGTC 2235
Qy 1561 TTCAGACTGAACCTGGGCCACTTTTACCTTGAAGCAGACAGAAAACATTATAGTAGCTAGCTTT 1620
Db 2236 TTCAGACTGAACCTGGGCCACTTTTACCTTGAAGCAGACAGAAAACATTATAGTAGCTAGCTTT 2295
Qy 1621 GATGGCAGAGGAAGTGGTTTACAAAGGAGATAAGATCATGCAATCAACAGAGACTG 1680
Db 2296 GATGGCAGAGGAAGTGGTTTACAAAGGAGATAAGATCATGCAATCAACAGAGACTG 2355

QY 1691 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 1740
Db GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 2415
QY 1741 GTGGACACAAACGAATTTGCAATTTGGGGCTGGTTCATATGAGGGTACGTAACCTCAATG 1800
Db GTGGACACAAACGAATTTGCAATTTGGGGCTGGTTCATATGAGGGTACGTAACCTCAATG 2475
QY 1801 GTCCTGGGATCGGGAAGTGGCGTGTTCAGTGTGGAAATAGCGCGCGCTGTATCCCGG 1860
Db GTCCTGGGATCGGGAAGTGGCGTGTTCAGTGTGGAAATAGCGCGCGCTGTATCCCGG 2535
QY 1861 TGGGAGTACTATGACTAGTGTACACAGAAAGTTACATGGGTCTCCCAACTCCAGAGAC 1920
Db TGGGAGTACTATGACTAGTGTACACAGAAAGTTACATGGGTCTCCCAACTCCAGAGAC 2595
QY 1921 AACCTTGACCAATACAGAAATTTCAACAGTCATGAGCAGCTGAAATTTTAAACAGTT 1980
Db AACCTTGACCAATACAGAAATTTCAACAGTCATGAGCAGCTGAAATTTTAAACAGTT 2655
QY 1981 GAGTACCTCTTATTCATGGAAACAGCAGATGATAACGTTCACTTTTCAGCAGTCACTCAG 2040
Db GAGTACCTCTTATTCATGGAAACAGCAGATGATAACGTTCACTTTTCAGCAGTCACTCAG 2715
QY 2041 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATACATGAA 2100
Db ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATACATGAA 2775
QY 2101 GACCATGGAATAGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160
Db GACCATGGAATAGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2835
QY 2161 ATAAAAACAATGTTCTCTTTTACCT 2184
Db ATAAAAACAATGTTCTCTTTTACCT 2859

RESULT 12

US-10-641-643-1026
; Sequence 1026, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Susan G. Stuart
; Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555

; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1026:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2924 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g180082
; SEQUENCE DESCRIPTION: SEQ ID NO: 1026 :
US-10-641-643-1026

Query Match 99.9%; Score 2180.8; DB 7; Length 2924;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2182; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGTCGCAAACTTACACTCTAACTGATTAATCTTAAAAAATACTTATAGACTGAAGTTATAC 60
Db AGTCGCAAACTTACACTCTAACTGATTAATCTTAAAAAATACTTATAGACTGAAGTTATAC 184
QY 61 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAAGAAAAATAATATCTTGGA 120
Db TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAAGAAAAATAATATCTTGGA 244
QY 121 TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATTTTGATGAGTTT 180
Db TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATTTTGATGAGTTT 304
QY 181 GGACATTTCTATCAATGATTAATTCATATCTCTGATGGCAGTTTATTTCTTTAGAAATAC 240
Db GGACATTTCTATCAATGATTAATTCATATCTCTGATGGCAGTTTATTTCTTTAGAAATAC 364
QY 241 AACTACGTGAAGCAATGGAGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300
Db AACTACGTGAAGCAATGGAGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 424
QY 301 AAAAGGCGAGCTGATTACAGAAGAGAGGATTTCCAAACAACAACACAGTGGGTCACTGTC 360
Db AAAAGGCGAGCTGATTACAGAAGAGAGGATTTCCAAACAACAACACAGTGGGTCACTGTC 484
QY 361 CCAGTGGGTCAATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACA 420
Db CCAGTGGGTCAATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACA 544
QY 421 AATTTACCAAGTTACAGAATCAATGACGCGGGAAGAAGATATAATATAAATGGAATA 480
Db AATTTACCAAGTTACAGAATCAATGACGCGGGAAGAAGATATAATATAAATGGAATA 604
QY 481 ACTGACTGGGTTTATGAAGAGGAGTCTTCAAGTGCCTACTCTGCTCTGTGTGGTCTCTCCA 540
Db ACTGACTGGGTTTATGAAGAGGAGTCTTCAAGTGCCTACTCTGCTCTGTGTGGTCTCTCCA 664
QY 541 AACGGCACTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db AACGGCACTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 724
QY 601 TCCTTCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db TCCTTCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 784
QY 661 GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTTGTGTAAATACAGACTCTCTCAGCTCA 720
Db GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTTGTGTAAATACAGACTCTCTCAGCTCA 844
QY 721 GTCAACCAATGAACTTCCATACAAATCACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db GTCAACCAATGAACTTCCATACAAATCACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 904
QY 781 TACTTGTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 840
Db TACTTGTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 964

QY 841 ATTGAGAACTATTCGGTTCATGGATATTTGTGACTATGATGAATCCAGTGAAGATGGAAC 900
Db
QY 965 ATTCAGAACTATTCGGTTCATGGATATTTGTGACTATGATGAATCCAGTGAAGATGGAAC 1024
Db
QY 901 TGCTTAGTGCGACGGCAACACATGTGAATGAGTACTGTGGCTGGGTTGGAAGATTTAGG 960
Db
QY 1025 TGCTTAGTGCGACGGCAACACATGTGAATGAGTACTGTGGCTGGGTTGGAAGATTTAGG 1084
QY 961 CCTTCAGAACCTCATTTTACCCCTTGATGGTAAATAGCTTCTACAGATCATCAGCAATGAA 1020
Db
QY 1085 CCTTCAGAACCTCATTTTACCCCTTGATGGTAAATAGCTTCTACAGATCATCAGCAATGAA 1144
QY 1021 GAAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCACATTTATTACA 1080
Db
QY 1145 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCACATTTATTACA 1204
QY 1081 AAGAGCACCTGGGAAGTCATCGGGATAGAAAGCTCTAAACAGTGAATTTATCTATCTACTACT 1140
Db
QY 1205 AAGAGCACCTGGGAAGTCATCGGGATAGAAAGCTCTAAACAGTGAATTTATCTATCTACTACT 1264
QY 1141 AGTAATGAAATATAAAGGAATGCCAGGAGGAGGAATCTTTATAAATCCAACTTATTGAC 1200
Db
QY 1265 AGTAATGAAATATAAAGGAATGCCAGGAGGAGGAATCTTTATAAATCCAACTTATTGAC 1324
QY 1201 TATACAAAAGTGACATGCCCTCAGTGTGAGCTGAATCCGGAAGAGTGTCAGTACTATTCT 1260
Db
QY 1325 TATACAAAAGTGACATGCCCTCAGTGTGAGCTGAATCCGGAAGAGTGTCAGTACTATTCT 1384
QY 1261 GTGTCATTCAGTAAAGGCGGAAGTATTTACGCTGAGATGTTCCGGTCTCGTCTGCC 1320
Db
QY 1385 GTGTCATTCAGTAAAGGCGGAAGTATTTACGCTGAGATGTTCCGGTCTCGTCTGCC 1444
QY 1321 CTCCTACTCTACAGCAGCGTGAATGATTAAGGCGCTGAGAGTCTCTGGAAGCAATTTCA 1380
Db
QY 1445 CTCCTACTCTACAGCAGCGTGAATGATTAAGGCGCTGAGAGTCTCTGGAAGCAATTTCA 1504
QY 1381 GCTTTGGATAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAAGCTGGACATTTATT 1440
Db
QY 1505 GCTTTGGATAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAAGCTGGACATTTATT 1564
QY 1441 TTGAATGAAACAAAAATTTTGGTATCAGATGATCTTGCCCTCCTCATTTTGTATAATCCAAG 1500
Db
QY 1565 TTGAATGAAACAAAAATTTTGGTATCAGATGATCTTGCCCTCCTCATTTTGTATAATCCAAG 1624
QY 1501 AATATCCTCTACTATTAGATGTGATGACAGGCCCATGTAGTCAAAAAGCAGACATGTC 1560
Db
QY 1625 AATATCCTCTACTATTAGATGTGATGACAGGCCCATGTAGTCAAAAAGCAGACATGTC 1684
QY 1561 TTGAGACTGAACTGGGCCCACTTACCTTGCAGCAGACAGAAAAATTTATAGTAGTACTGTT 1620
Db
QY 1685 TTGAGACTGAACTGGGCCCACTTACCTTGCAGCAGACAGAAAAATTTATAGTAGTACTGTT 1744
QY 1621 GATGGCAGAGGAAGTGGTTACCAAGAGATAAGATCATGCTCAATCAACAGAGACTG 1680
Db
QY 1745 GATGGCAGAGGAAGTGGTTACCAAGAGATAAGATCATGCTCAATCAACAGAGACTG 1804
QY 1681 GGAACATTTGAAAGTTCGAAGTCAAAATTTGAAGCGCCAGACAAATTTCAAAAATGGGATTT 1740
Db
QY 1805 GGAACATTTGAAAGTTCGAAGTCAAAATTTGAAGCGCCAGACAAATTTCAAAAATGGGATTT 1864
QY 1741 GTGGACAAACAAACGAATTTGCAATTTTGGGGCTGGTCAATATGGAGGGTACGTTAACTCAATG 1800
Db
QY 1865 GTGGACAAACAAACGAATTTGCAATTTTGGGGCTGGTCAATATGGAGGGTACGTTAACTCAATG 1924
QY 1801 GTCCCTGGGATCGGGAAAGTGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1860
Db
QY 1925 GTCCCTGGGATCGGGAAAGTGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1984
QY 1861 TGGGAGTACTATGACTCAGTGTATACAGAAACGTTATCATGGGTCTCCAACTCCAGAGAC 1920
Db
QY 1985 TGGGAGTACTATGACTCAGTGTATACAGAAACGTTATCATGGGTCTCCAACTCCAGAGAC 2044
QY 1921 AACCTTGACCATTTACAGAAATTTCAACAGTCATGAGCAGAGCTGAAAAATTTTAAACAAAGTT 1980

Db 2045 AACCTTGACCATTTACAGAAATTTCAACAGTCATGAGCAGAGCTGAAAAATTTTAAACAAAGTT 2104
QY 1981 GAGTACCTCTCTTATTCATGGAACACAGCAGATGATAGTTCCTTTCCAGCAGTCAGCTCAG 2040
Db 2105 GAGTACCTCTCTTATTCATGGAACACAGCAGATGATAGTTCCTTTCCAGCAGTCAGCTCAG 2164
QY 2041 ATCTCCAAAGCCCTCGTGCATGTTGGAGTGTTCAGGCAATGTGGTATATCTGATGAA 2100
Db 2165 ATCTCCAAAGCCCTCGTGCATGTTGGAGTGTTCAGGCAATGTGGTATATCTGATGAA 2224
QY 2101 GACCATGGAATAGCTAGCAGCAGCAGCACCAACATATATATACCCACATGAGCCACTTC 2160
Db 2225 GACCATGGAATAGCTAGCAGCAGCAGCACCAACATATATATACCCACATGAGCCACTTC 2284
QY 2161 ATAAAAACAATGTTTCTCTTTTACCT 2184
Db 2285 ATAAAAACAATGTTTCTCTTTTACCT 2308

RESULT 13
US-10-631-467-969
; Sequence 969, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive l
; TITLE OF INVENTION: disease
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 969
; LENGTH: 3316
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-631-467-969

Query Match 76.8%; Score 1677.6; DB 9; Length 3316;
Best Local Similarity 86.0%; Pred. No. 0;
Matches 1876; Conservative 0; Mismatches 294; Indels 12; Gaps 1;

QY 1 AGTCGCAAAAACCTTACACTCTAACTGATTACTTTAAAAAATACTTATAGACTGAAGTTATAC 60
Db 226 AGCGCAGAACGATTTCACTAGCTGACTATTTAAAGAGTACCTTTCCGGTCAAGTCCCTAC 285
QY 61 TCCTTAAAGATGGATTTTCAGATCATGAATATCTCTCAAAACAAGAAAAATATATCTTGGTA 120
Db 286 TCTTTGTGGTGGTTTCAGACTTTTGAATACCTCTACAAACAAGAACAAATATCTTGTCTG 345
QY 121 TTCATGCTGATATGGAACAGCTCAGTCTTTCTTGGAGAACAGTACATTTGATGATGTT 180
Db 346 CTCATGCTGAACATGGAACAGCTCCATTTCTTGGAGAACAGTACCTTTGAAAGCTTT 405
QY 181 GGACATTTCTATCAATGATTATTTCAATATCTCTGATGGCAGTTTATTTCTTTAGAAATAC 240
Db 406 GGA-----TATCATTTAGTGTCACTGACCCAGCTGTTTCTCTTGGAAATAC 453
QY 241 AACTAGCTGAAGCAATGGAGCATTTCTTACACAGCTTCATATGACATTTATGATTTAAAT 300
Db 454 AACTAGCTGAAGCAATGGAGCATTTCTTACACAGCTTCATATCAACATTTATGATGTGAAT 513
QY 301 AAAAGCAGCTGATTTACAGAGAGAGGATTTCCAAACACACACAGCTGGTGCATGTCATCA 360
Db 514 AAAAGCAGCTGATTTACAGAGAGAGGATTTCCAAATAATACACAGTGGATTCATGGTCA 573
QY 361 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACATGATTTATGTTTAAATTTGAACCA 420


```
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1570
; LENGTH: 4835
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012789
US-09-917-800A-1570

Query Match          75.8%; Score 1656.2; DB 3; Length 4835;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 1863; Conservative 0; Mismatches 313; Indels 9; Gaps 1;

QY      7  AAAAATTACACTCTAACTGATTACTTTAAAAAATACCTTATAGACTGGAAGTTATACCTCTTA 66
Db      203 AGAACTTTACACTAGCTAGCTATATTTAAAGATACCTTTGGGTCAGTCTCTACTCTCTTG 262

QY      67  AGATGGATTTTCAGATCATGAATATCTCTACAAACGAAGAAAATAATATCTTGGTATTCAAT 126
Db      263 CGGTGGGTTTCAGATTTCTGAATACCTCTACAAAGCAAGAAAACAATATCTTGTCTATTCAAT 322

QY      127 GCTGAATATGGAACACAGCTCAGTTTTCTTGGAGAACAGTACATATTTGATGAGTTTGGACAT 186
Db      323 GCTGAACACGGGAACAGCTCCATTTTCTTGGAGAACAGTACCTTTTGGAGATCTTTGGAGAT 382

QY      187 TCTATCAATGATTATTTCATATCTCTGATGGGAGTATTATCTTTAGAAATACAACTAC 246
Db      383 TCTATNAAGTATTTTCACTGTGTCACCCGACAGACTGTTGCTTCTTTAGAAATACAAATAT 442

QY      247 GTGAAGCAATGGAGGCAITTCCTACACAGCTTCATATGACATTTATGATTTAAATAAAAGG 306
Db      443 GTGAAGCAATGGAGACACTCTACACGGCTTCATACAGTATTTATGACTTTGAATAAAAGA 502

QY      307 CAGCTGATTTACAGAGAGAGGATTCAAAACACACACAGTGGGTACATGGTCCACAGTG 366
Db      503 CAGCTGATTCACAGAGAGAGGATTCAAAATAATACACAGTGGATCACATGGTCCACAAGAA 562

QY      367 GGTCAATAATTCGCATATGTTGGAAACAATGACATTTATGTTAAATTCGAACCAATTTA 426
Db      563 GGTCAATAATTCGCATATGTTGGAAAGAATGATATTTATGTTAAATTTGAAACCAATTTG 622

QY      427 CCAAGTTACAGAAATCACATGGACGGGGAAGAGATATAATATATAATGAAATAAATGAC 486
Db      623 CCTAGTCATAGGATCACATCAACAGGAAAGAAAATGTAATATTTAAACGGAATAAATGAC 682

QY      487 TGGGTTTATGAAGAGAAAGTCTTCAGTGCCTACTCTGCTCTCTGTTGGTGTCTCAACAGGC 546
Db      683 TGGGTTTATGAAGAGAAATCTTCGGTGCCTACTCCGCACACTGTGGTGGTCTCCAAACGGC 742

QY      547 ACTTTTTTAGCATATGCCCCAATTTAACGACACAGAGTCCCACTTTATGCAATACCTCTTC 606
Db      743 ACTTTTCTAGCTTATGCCCCAGTTTAAACGACACGGAGTGCCTCTCATTTGAATACTCTCTC 802

QY      607 TACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCCATATATCAAAAGGCAGGA 666
Db      803 TACTCTGATGAGTCACTGAGTACCCCAAGACAGTCTGGAATCCGTTACCCAAAGGCAGGA 862

QY      667 GCTGGAATCCAACTGTAAGTTCTTTTGTGTAATAATACAGACTCTCTCAGCTCAGTCAACC 726
Db      863 GCTGTGAATCCAACTGTAAGTTCTTTTATTTGTAATAATACAGACTCTCTCAGCTCAACTACT 922

QY      727 AATGCAACTTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGGATCACTACTTTG 786
```

```
Db      923 ACTACGATTCCTCATGCAAAATCACCGCTCCTGCTATCTGTGACAAACAGGGGATCACTACTTG 982
QY      787 TGTGATGTGACATGGGCAACACAGAAAGAATTTCTTTGTCAGTGGCTCAGGAGGATTCAAG 846
Db      983 TGTGACGTGGCTGGGTTTTCAGAAAGACAGAAATCTCGTTGTCAGTGGCTCAGGAGGATTCAAG 1042
QY      847 AACTATTCCGTCATGATATTTTGTGACTATATGATGAATCCAGTGGAGAGATGGAACCTGCTTA 906
Db      1043 AACTATTCCGTCATGATGCGGATCTGCGACTATGATAGACACACCTAGTATGGAACCTGCTCA 1102
QY      907 GTGGCACGGCAACACATTTGAAATGAGTACTACTTGGCTGGGTTTGGAGAGATTTTAGGCTTCA 966
Db      1103 ACGACGACAGGAGCATATTTGAAACAGTGCACAGGCTGGTGGGAGAGATTTTAGGCTGCA 1162
QY      967 GAACCTCATTTTACCTTTGATGGTATAGCTTCTTCAAGATCATCAGCAATGAAGAGT 1026
Db      1163 GAACCCCACTTCACTCCGACGGAAGCAGCTTCTATAAATCTGTCAGTGAACAAGATGCG 1222
QY      1027 TACGACACATTTGCTATTTTCCAAATAGATA-----AAAAAGACTGCACATTTAT 1077
Db      1223 TACAACACATCTGCCAGTTCAGAAAGATAGAAACCCGAAACAGGTCGTACATTTAT 1282
QY      1078 ACAAAAGGCACCTGGGAAAGTCACTCGGATAGAAGCTCTAACCAAGTGAATTTATATACTATAC 1137
Db      1283 ACANAAGGAGCTGGGAAAGTCACTTAGTATCGAAGCTCTGACCAGCGATTTATCTGTACTAC 1342
QY      1138 ATTAGTAATGAATATAAAAGGAATGCCAGGAGGAAGGAATCTTTATAAATCCAACTTAT 1197
Db      1343 ATTAGTAATGAATATAAAAGGAATGCCAGGAGGAAGGAATCTTTATAAATTTCACTTACT 1402
QY      1198 GACTATACAAAGTGCACATGCTCAGTTGTGAGCTGAAATCCGAAAGGTGTGAGTACTAT 1257
Db      1403 GACCAACAATATAAGAGTGCCTTAGTTGTGACCTGAAATCCAGAAAGATGCCAGTATATAC 1462
QY      1258 TCTGTGTCAITTCAGTAAAGAGCGGAAGTATTTATCAAGTGAAGTGTTCGGGTCTCTGGTCTG 1317
Db      1463 TCGGTGTCACTTAGTAAAGAGCAAGTACTATCAGCTGGGATCCGGGGCCCTGGTCTG 1522
QY      1318 CCCCTCTATCTTACACAGCAGCGTGAATGATTAAGGGCTCAGAGTCTCTGGAAGACAT 1377
Db      1523 CCCCTCTACACTCTGCACTCGCAGCACTGATCAAAAAGAGCTGAGAGTCTCTGGAGGACAT 1582
QY      1378 TCAGCTTTGGATATAAATGCTGCAGAACTGTCAGATGCTCCAGATGCCCTCAAAAACCTGGACTTCA 1437
Db      1583 TCTGTTTGGATATAAATGCTGCAGATGCTCMAATGCTTCAAAAATTTGACTTCAAT 1642
QY      1438 ATTTTGAATGAAACAAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGAATAATCC 1497
Db      1643 GTTCTGATGAAACAAAGATTTTGGTATCAAAATGATCTTTACCTCTCTCATTTTGAATAATCC 1702
QY      1498 AAAAAATATCTCTACTATTTAGATGTGTATGACGAGCCCATGTAGTCAAAAAGCAGACAT 1557
Db      1703 AAAAAATATCTCTACTATTTAGATGTATATGCAAGTCTCTGTAGTCAAAAAGCAGATGCT 1762
QY      1558 GTCTTCAGACTGAATGGGCGCACTTACCTTGAACACAGAAACATATATAGTAGTAGC 1617
Db      1763 GCCTTCAGACTGAATGGGCGCACTTACCTTGAACACAGAAACATATAGTAGTAGC 1822
QY      1618 TTTGATGACAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATCAACGAAGA 1677
Db      1823 TTTGATGACAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATCAACGAAGA 1882
QY      1678 CTGGAAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGGCGAGCAATTTTCAAAAATGGGA 1737
Db      1883 CTGGAAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGGCGAGCAATTTTCAAAAATGGGA 1942
QY      1738 TTTTGGGACAAACAAAGATTTGCAATTTGGGCTGCTCATATGAGGGGTACGTAACTCA 1797
Db      1943 TTTTGGGACAGCAACGAGTGTGCAATTTGGGGCTGGTCTATATGAGGGGTACGTAACTCA 2002
QY      1798 ATGTGCTCTGGGATCGGGAAAGTGGCGTGTTCAGTGTGGAATAGCCGTGGCGCTGTATCC 1857
```


Db 2003 ATGTCCTGGGANTCGGGAAGTGGCGTGTTCAAGTGTGGAATAGCCGTGGCGCCCGTGCA 2062
Qy CGGTGGGAGTACTATGACTCAGTGTACACAGAACGTTATACATGGGTCTCCCACTCCAGAA 1917
Db CGGTGGGAGTACTATGACTCAGTGTACACAGAACGTTATACATGGGTCTCCCACTCCAGAG 2122
Qy GACAACTTGACCATTAACAGAAATTCACAGTCAATGACAGAGCTGAAATTTTAAACAA 1977
Db GACAACTTGACCATTAACAGAAATTCACAGTCAATGACAGAGCTGAAATTTTAAACAA 2182
Qy GTTCAGTACCTTCCTTATTCATGGAACACAGACAGATGATACGTTCACTTTACAGAGTCAGCT 2037
Db GTTCAGTACCTTCCTTATTCACGGGTACACAGATGATGATGTTCACTTTACAGAGTCAGCT 2242
Qy CAGATCTCAAAAGCCCTCGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATACTGAT 2097
Db CAGATCTCAAAAGCCCTCGTGGATGCTGGCGTGGATTTCCAGGCAATGTGGTATACGGAC 2302
Qy GAAGACCATGGAATAGCTAGCAGCAGCAGCAACCAACATATATATACCAATGAGCCAC 2157
Db GAAGACCATGGAATGCGCAGCAGCAGCAGCTCACAGCACATCTATTCCACATGAGCCAT 2362
Qy TTCATAAACAAATGTTCTCTTTAC 2182
Db TTCTCCAGCAGTGCTTCTCTTAC 2387

RESULT 15

US-10-165-603-5
; Sequence 5, Application US/10165603
; Publication No. US20030021792A1
; GENERAL INFORMATION:
; APPLICANT: Roben, Paul W.
; APPLICANT: Stevens, Anthony C.
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: TPTECH.001A
; CURRENT APPLICATION NUMBER: US/10/165,603
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/297,021
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/305,117
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4835
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-165-603-5

Query Match 75.8%; Score 1656.2; DB 5; Length 4835;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 1863; Conservative 0; Mismatches 313; Indels 9; Gaps 1;

Qy 7 AAAACTTACACTCTAACTGATTAATAAATACTTATAGACTGAAGTTATCTCTTCA 66
Db AGAATTACACACTAGTGACTATTTAAAGATACCTTTCGGGTCAAGTCTCTCTTG 262
Qy 67 AGATGGATTTGAGATCATGATATCTCTACAAACAGAAATAATATCTTGGTATTCAT 126
Db CGGTGGGTTTCAGATTTCTGAATACCTCTACAGCAAGAAACAATATCTTCTGATTCAT 322
Qy 127 GCTGAATAGGAACAGCTCAGTTTCTTGGAGACAGTACATTTGATGATTTGGACAT 186
Db GCTGAACACGGGAACAGCTCCATTTTCTTGGAGAACAGTACCTTTGAGATCTTTGGAGAT 382
Qy 187 TCTATCAATGATTTCAATATCTCTGATGGCAGTTTATCTCTTAGAATACAACTAC 246
Db TCTATAGTATTTTCAAGTCTCAGTCTCAGCAGACATGTTCTTCTTAGAATACAAATAT 442
Qy 247 GTGAAGCAATGGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAATAAAGG 306

Db 443 GTGAAGCAATGGAGGACACTCTCTACACGGCTTCATACAGTATTTATGACTTGAATAAAGA 502
Qy CAGCTGATTTACAGAAAGAGAGGATTTCCAAACAAACACACAGTGGGTCACTAGTGCACAGTG 366
Db CAGCTGATTTACAGAAAGAGAGGATTTCCAAATATATACACAGTGGATCATAGTGCACAGAA 562
Qy GGTCAATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCAATTTA 426
Db GGTCAATAAATTTGGCATATGTTCTGGAAGATGATATTTATGTTAAATTTGAACCAATTTG 622
Qy CCAAGTTACAGAAATCACATGACCGGGGAGAGAAATATATATATATATATGGAATTAAGTAC 486
Db CCTAGTCATAGGATCACATCAACAGGAGAAAGAAATGTAATATTTAAACGGAATTAATGAC 682
Qy TGGGTTTATGAAGAGAGGAGTCTTCAGTGCCTACTCTGCTGTGTGTGTCTCTCAACCGC 546
Db TGGGTTTATGAAGAGAGAAATCTTCGGTGCCTACTCTCGCACATGTGTGTGTCTCTCAACCGC 742
Qy ACTTTTATGATATGCCCCAATTTAACGACACAGAAAGTCCCACTTATTTGAATATCTCTTTC 606
Db ACTTTTCTAGCTTATGCCCCAGTTTAAACGACACCGAGTGCCTCTCATTTGAATATCTCTTTC 802
Qy TACTCTGATGAGTCACTGCGAGTACCCAAAGACTGTACGGGTTCATATCCAAAGGACAGA 666
Db TACTCTGATGAGTCACTGCGAGTACCCCAAGACAGTCTGTGGATTCGGTACCCAAAGGACAGA 862
Qy GCTGTGAATCCAACTGTAAAGTCTTTGTTGTAATACAGACTCTCTCAGCTCAGTCACT 726
Db GCTGTGAATCCAACTGTAAAGTCTTTGTTGTAATACAGACTCTCTCAGCTCAGTCACT 922
Qy AATGCAACTTCCATCAAAATCACTGCTCTGCTCTTCTATGTTGATAGGGGATCACTACTTG 786
Db ACTACGATTTCCATGCAAAATCACCGCTCTCTGATCTGTGACACACAGGGGATCACTACTTG 982
Qy TGTGATGATGAGTGGGCAACACAGAAAGAAATTTCTTTCAGTGGCTCAGAGGATTCAG 846
Db TGTGACGTGGCTGGGTTTCAGAAAGACAGAAATCTCGTTGACGTGGCTCAGAGGATTCAG 1042
Qy AACTATTCCGGTCATGGATATTTGTGACTATGATGAATCCAGTGGAAAGATGGAATGCTTA 906
Db AACTATTCCGGTCATGGATATTTGTGACTATGATGAATGGAATGGAATGGAATGCTTA 1102
Qy GTGCAACGGCAACACATTTGAAATCAGTACTACTGCTGGGTGGTGGAAAGATTTAGGCTTCA 966
Db ACAGCAGCAGGAGCATATTTGAAACAGTGCACAGGCTGGTGGCGGAAGATTTAGGCTTCA 1162
Qy GAACCTCATTTTACCTTGATGTTAGTCTTACAAAGATCATCAGCAATGAGAAAGT 1026
Db GAACCCCACTTTCACCTCCGACGGAAAGCAGCTTCTATAAAATCGTCAGTGACAAAGATGGC 1222
Qy TACAGACACATTTGCTATTTTCCAAATAGATA-----AAAAAGACTGCACATTTATT 1077
Db TACAAACACATCTCCAGTTCAGAAAGATAGGAAACCCGAAACAGGCTCTGTACATTTATT 1282
Qy ACAAAGGACCTCGGGAAGTCACTAGTATCGAAGCTCTGACCGAGATTTATCTGTACTAC 1137
Db ACAAAGGACCTCGGGAAGTCACTAGTATCGAAGCTCTGACCGAGATTTATCTGTACTAC 1342
Qy ATTAGTATGAAATTAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATT 1197
Db ATTAGTATGAAATTAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATT 1402
Qy GACTATACAAAGTGACATGCTCAGTTGTGAGTGAATCCGAAAGGTGTCAGTACTAT 1257
Db GACCAACAATTAAGAAAGTGCCTTAGTTGTGACCTGAAATCCAGAAAGATGCAAGTATTAC 1462
Qy TCTGTGCTATTTCAGTAAAGGCGGAAGTATTTATCAGCTGAGATGTTCCGGTCTGTGCTG 1317
Db TCGGTGCTACTTAGTAAAGGCGCAAGTACTATCAGCTGGGATGCCGGGGCCCTGGTCTG 1522
Qy CCCCTCTATCTTACACAGCAGCGGTGAATGATAAAGGGCTGAGAGTCTCGGAAGACAAAT 1377
Db CCCCTCTACACTCTGCAATCGCAGCACTGATCAAAAGAGAGCTGAGAGTCTCTGGAGGACAAAT 1582

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2006, 04:36:39 ; Search time 400 Seconds
(without alignments)
9705.487 Million cell updates/sec

Title: US-10-659-055-2
Perfect score: 2184
Sequence: 1 agtgcgaaaacttacctct.....aacaaatgtttcttacct 2184

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2184	100.0	3407	3	US-10-002-593-5
2	2184	100.0	3407	3	US-09-949-016-275
3	2184	100.0	3407	3	US-10-423-714-5
4	2182.4	99.9	3419	3	US-09-949-016-4579
5	2180.8	99.9	2924	3	US-09-021-655-1026
6	2180.8	99.9	2924	6	PCT-US93-07923-1
7	751	34.4	2815	2	US-08-230-491A-1
8	751	34.4	2815	2	US-08-619-280A-1
9	751	34.4	2815	2	US-08-940-391-1
10	751	34.4	2815	3	US-09-265-606-1
11	669.4	30.7	2736	3	US-09-949-016-2459
12	264.8	12.1	535	3	US-09-389-681-428
13	264.8	12.1	535	3	US-09-620-405B-428
14	264.8	12.1	535	3	US-09-433-826B-428
15	264.8	12.1	535	3	US-09-604-287A-428
16	264.8	12.1	535	3	US-09-834-759-428
17	264.8	12.1	535	3	US-09-590-751A-428
18	264.8	12.1	535	3	US-09-551-621-428
19	264.8	12.1	535	3	US-09-551-621A-428
20	264.8	12.1	535	3	US-10-076-622-428
21	215.8	9.9	2576	3	US-10-363-937-21
22	215.8	9.9	2583	3	US-09-976-674-6
23	215.8	9.9	4541	3	US-09-976-674-42
24	197	9.0	85869	3	US-09-949-016-12017

25	197	9.0	85878	3	US-09-949-016-16321	Sequence 16321, A
26	177.4	8.1	4496	3	US-09-976-674-44	Sequence 44, Appl
c 27	141	6.5	239	3	US-09-392-184-32	Sequence 32, Appl
28	121.8	5.6	1023	3	US-09-270-767-13726	Sequence 13726, A
c 29	109.8	5.0	601	3	US-09-949-016-23690	Sequence 23690, A
c 30	109.8	5.0	601	3	US-09-949-016-163433	Sequence 163433, A
c 31	99	4.5	657	3	US-09-221-017B-646	Sequence 646, App
32	92.6	4.2	4280	3	US-09-079-592-1	Sequence 1, Appl
33	92.6	4.2	4280	3	US-09-462-284-1	Sequence 1, Appl
34	88	4.0	1485	3	US-09-248-796A-5369	Sequence 5369, Ap
c 35	88	4.0	4982	3	US-08-699-103B-1	Sequence 1, Appl
c 36	88	4.0	4982	3	US-09-223-059-1	Sequence 1, Appl
c 37	88	4.0	4982	3	US-09-628-133-1	Sequence 1, Appl
38	81	3.7	547	3	US-09-280-116-249	Sequence 249, App
39	77.8	3.6	2671	3	US-09-976-674-2	Sequence 2, Appl
40	77.8	3.6	2797	3	US-09-976-594-1103	Sequence 1103, Ap
41	77.8	3.6	3120	3	US-10-070-464-2	Sequence 2, Appl
42	77.8	3.6	4309	3	US-09-976-674-14	Sequence 14, Appl
43	77.8	3.6	4685	3	US-09-976-674-22	Sequence 22, Appl
44	77.8	3.6	4829	3	US-09-976-674-12	Sequence 12, Appl
45	76.6	3.5	76848	3	US-09-949-016-14201	Sequence 14201, A

ALIGNMENTS

RESULT 1
US-10-002-593-5
; Sequence 5, Application US/10002593
; Patent No. 6586198
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERT
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Atty Docket No. 6586198 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002.593
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 3407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-002-593-5

Query Match	100.0%	Score 2184;	DB 3;	Length 3407;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2184;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGTCGAAAACCTTACACTCTTAACCTGATTACTTAAAAAATATTATAGACTGAACTTATAC	60	
Db	190	AGTCGAAAACCTTACACTCTTAACCTGATTACTTAAAAAATATTATAGACTGAACTTATAC	249	
Qy	61	TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAAGAAATAAATATCTTGTA	120	
Db	250	TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAAGAAATAAATATCTTGTA	309	
Qy	121	TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATCTTGGAGTTTT	180	
Db	310	TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATCTTGGAGTTTT	369	
Qy	181	GGACATTTCTATCAATGATTAATTCATATCTCTGATGGGCAAGTTATCTCTTTAGAAATAC	240	
Db	370	GGACATTTCTATCAATGATTAATTCATATCTCTGATGGGCAAGTTATCTCTTTAGAAATAC	429	
Qy	241	AACCTACCTGAGCAATGGAGCATTCCTACACAGCTTCATATGACATTTATGATTTAAAT	300	
Db	430	AACCTACCTGAGCAATGGAGCATTCCTACACAGCTTCATATGACATTTATGATTTAAAT	489	
Qy	301	AAAAGGCGAGCTGATTACAGAAAGAGAGGATTCCTCAAAACAACACACAGTGGGTCACTGGTCA	360	

Db 1750 TTGAGCTGAACTGGGCCACTTACCTTGAAGCAGAGAAACATTTATAGTAGCTAGCTTT 1809
Qy 1621 GATGCGAGAGAGTGGTTTACCAAGGAGATAGATCATGTCATCAATCAACAGAGAGCTG 1680
Db 1810 GATGCGAGAGAGTGGTTTACCAAGGAGATAGATCATGTCATCAATCAACAGAGAGCTG 1869
Qy 1681 GGAACTTTGAAGTTGAAGATCAAAATGAAGCAGCCAGACAATTTTCAAAAATGGGATTT 1740
Db 1870 GGAACTTTGAAGTTGAAGATCAAAATGAAGCAGCCAGACAATTTTCAAAAATGGGATTT 1929
Qy 1741 GTGGACAAACAAAGTTCGAATTTGGGCTGGTTCATATGAGGGTACGTAACCTCAATG 1800
Db 1930 GTGGACAAACAAAGTTCGAATTTGGGCTGGTTCATATGAGGGTACGTAACCTCAATG 1989
Qy 1801 GTCTGGGATCGGGAAGTGGGGTGTTCAGTGTGGAATAGCGGTGGCGCTGTATCCCGG 1860
Db 1990 GTCTGGGATCGGGAAGTGGGGTGTTCAGTGTGGAATAGCGGTGGCGCTGTATCCCGG 2049
Qy 1861 TGGAGTACTATGACTCAGTGTACACAGAACTGTTACATGGGTCTCCAACTCCAGAGAC 1920
Db 2050 TGGAGTACTATGACTCAGTGTACACAGAACTGTTACATGGGTCTCCAACTCCAGAGAC 2109
Qy 1921 AACCTTGACCAATTACAGAAATTCACAGTTCATGAGCAGCTGGAATTTTAAACAGTT 1980
Db 2110 AACCTTGACCAATTACAGAAATTCACAGTTCATGAGCAGCTGGAATTTTAAACAGTT 2169
Qy 1981 GAGTACCTCTTATTCATGGAACAGCAGATGATAACGTTTCACTTTCAGCAGTCACTCAG 2040
Db 2170 GAGTACCTCTTATTCATGGAACAGCAGATGATAACGTTTCACTTTCAGCAGTCACTCAG 2229
Qy 2041 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATACTGATGAA 2100
Db 2230 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATACTGATGAA 2289
Qy 2101 GACCATGGAATAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160
Db 2290 GACCATGGAATAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2349
Qy 2161 ATAAACAAATGTTCTCTTTACCT 2184
Db 2350 ATAAACAAATGTTCTCTTTACCT 2373

RESULT 4
US-949-016-4579
; Sequence 4579, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4579
; LENGTH: 3419
; TYPE: DNA
; ORGANISM: Human
US-949-016-4579
Query Match 99.9%; Score 2182.4; DB 3; Length 3419;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGTCGCAAACTTACACTCTAACTGATTACTTTAAATAAATACTTATAGACTGAAGTTATAC 60

Db 191 AGTCGCAAACTTACACTCTAACTGATTACTTTAAAAATACTTATAGACTGAAGTTATAC 250
Qy 61 TCCTTATAGATGGATTTAGATCATGATATCTCTACAAACAAGAAAATAATATCTTTGGTA 120
Db 251 TCCTTATAGATGGATTTAGATCATGATATCTCTACAAACAAGAAAATAATATCTTTGGTA 310
Qy 121 TTCAATGCTGAATATGGAAACAGCTCAGTTTTCTTTGGAGAACAGTACATTTTATGATGAGTTT 180
Db 311 TTCAATGCTGAATATGGAAACAGCTCAGTTTTCTTTGGAGAACAGTACATTTTATGATGAGTTT 370
Qy 181 GGAATTTCTATCAATGATTTATTTCAATATCTCTGATGGCAGTTTATTTCTTTAGAAATAC 240
Db 371 GGAATTTCTATCAATGATTTATTTCAATATCTCTGATGGCAGTTTATTTCTTTAGAAATAC 430
Qy 241 AACTACGTTAGCAATGGAGCATTTCTTACACAGCTTCATATGACATTTATGATTTAAT 300
Db 431 AACTACGTTAGCAATGGAGCATTTCTTACACAGCTTCATATGACATTTATGATTTAAT 490
Qy 301 AAAAGGCGAGCTGATTTACAGAGAGAGGATTTCCAAACAACACACAGTGGGTCACTGGTCA 360
Db 491 AAAAGGCGAGCTGATTTACAGAGAGAGGATTTCCAAACAACACACAGTGGGTCACTGGTCA 550
Qy 361 CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATTTAAATTTGAAACA 420
Db 551 CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATTTAAATTTGAAACA 610
Qy 421 AATTTTCCAAAGTTACAGATTCACATGAGCGGGAAGAGATATATATATATATATGATTA 480
Db 611 AATTTTCCAAAGTTACAGATTCACATGAGCGGGAAGAGATATATATATATATATGATTA 670
Qy 481 ACTGACTGGGTATTAAGAGAGAGTCTTTCAGTGCCTACTCTGCTCTGCTGGTGGTCTCCA 540
Db 671 ACTGACTGGGTATTAAGAGAGAGTCTTTCAGTGCCTACTCTGCTCTGCTGGTGGTCTCCA 730
Qy 541 AACGGCACTTTTTPAGCATATGCCCAATTTAAACGACACAGAACTCCCACTTTATTTGAAAT 600
Db 731 AACGGCACTTTTTPAGCATATGCCCAATTTAAACGACACAGAACTCCCACTTTATTTGAAAT 790
Qy 601 TCCTTCTACTCTGATGAGTCACTGACGATCCCAAGAGCTGACGGGTTCATATCCCAAG 660
Db 791 TCCTTCTACTCTGATGAGTCACTGACGATCCCAAGAGCTGACGGGTTCATATCCCAAG 850
Qy 661 GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTTGTGTAAATACAGACTCTCTCAGCTCA 720
Db 851 GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTTGTGTAAATACAGACTCTCTCAGCTCA 910
Qy 721 GTCACCAATGCAACTTCCATACAAATCACTGCTCTCTCTATGTTGATAGGGGATCAC 780
Db 911 GTCACCAATGCAACTTCCATACAAATCACTGCTCTCTCTATGTTGATAGGGGATCAC 970
Qy 781 TACTTGTGATGTGATGATGGGCAACACAGAAAGAAATTTCTTTGAGTGGCTCAGGAGG 840
Db 971 TACTTGTGATGTGATGATGGGCAACACAGAAAGAAATTTCTTTGAGTGGCTCAGGAGG 1030
Qy 841 ATTCCAGAACTATTTCGGTTCATGATATTTGATGACTATGATGATGATGATGATGATGAT 900
Db 1031 ATTCCAGAACTATTTCGGTTCATGATATTTGATGACTATGATGATGATGATGATGATGAT 1090
Qy 901 TGCTTAGTGCAAGGCAACACATTTGAAATCAGTACTACTGCTGGTGGTGGAAAGATTTAGG 960
Db 1091 TGCTTAGTGCAAGGCAACACATTTGAAATCAGTACTACTGCTGGTGGTGGAAAGATTTAGG 1150
Qy 961 CCTTCCAGAACTCAATTTTACCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 1151 CCTTCCAGAACTCAATTTTACCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1210
Qy 1021 GAAGGTTACAGACACATTTTCTTATTTCCAAATAGATAAAAAAGACTGCACATTTATTTACA 1080
Db 1211 GAAGGTTACAGACACATTTTCTTATTTCCAAATAGATAAAAAAGACTGCACATTTATTTACA 1270
Qy 1081 AAAGGCACTGGGAGAGTCACTGGGATAGAAAGCTCTTAAACAGTATTATCTATCTACTACATT 1140

Db 1271 AAAGGCACCTGGAGTGCATCGGATAGAGCTTAAACAGTGATTTATCTATCTACTACATT 1330
Qy 1141 AGTAATGAATATAAGGAATGCCAGGAGGAATCTTTTATAAATCAACTTATTGAC 1200
Db 1331 AGTAATGAATATAAGGAATGCCAGGAGGAATCTTTTATAAATCAACTTATTGAC 1390
Qy 1201 TATACAAAAGTCACATGCTCAGTTGTGTAGCTGAATCCGAAAGGTGTGAGTACTATTCT 1260
Db 1391 TATACAAAAGTCACATGCTCAGTTGTGTAGCTGAATCCGAAAGGTGTGAGTACTATTCT 1450
Qy 1261 GTGTCAATTCAGTAAAGAGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCGTCTGCC 1320
Db 1451 GTGTCAATTCAGTAAAGAGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCGTCTGCC 1510
Qy 1321 CTCCTATCTACACAGCAGCGTGAATGATTAAGGCTGAGAGTCTCGAAGACAAATTC 1380
Db 1511 CTCCTATCTACACAGCAGCGTGAATGATTAAGGCTGAGAGTCTCGAAGACAAATTC 1570
Qy 1381 GCTTTGGATAAAATGCTGCAGATGTCAGATGCCCTCCAAAATACTGGACTTCATTATT 1440
Db 1571 GCTTTGGATAAAATGCTGCAGATGTCAGATGCCCTCCAAAATACTGGACTTCATTATT 1630
Qy 1441 TTGAATGAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAAATCC 1500
Db 1631 TTGAATGAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAAATCC 1690
Qy 1501 AAATATCCTCTACTATTAGATGTATGAGGCGCCATGTAGTCAAAAAGCAGACACTGTC 1560
Db 1691 AAATATCCTCTACTATTAGATGTATGAGGCGCCATGTAGTCAAAAAGCAGACACTGTC 1750
Qy 1561 TTCAAGCTCACTGGGCACTTACCTTGAAGCAGACAGAAAATATATAGTACTAGCTTT 1620
Db 1751 TTCAAGCTCACTGGGCACTTACCTTGAAGCAGACAGAAAATATATAGTACTAGCTTT 1810
Qy 1621 GATGCGAGAGGAGTGGTTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGACTG 1680
Db 1811 GATGCGAGAGGAGTGGTTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGACTG 1870
Qy 1681 GGAACATTTGAAGTTGAAGATCAATTTGAAGCAGCAGACAATTTTCAAAAATGGGATTT 1740
Db 1871 GGAACATTTGAAGTTGAAGATCAATTTGAAGCAGCAGACAATTTTCAAAAATGGGATTT 1930
Qy 1741 GTGGACCAACAAAGTAATTCGAATTTGGGGCTGGTCAATGAGGAGTACGTAACCTCAATG 1800
Db 1931 GTGGACCAACAAAGTAATTTGGGGCTGGTCAATGAGGAGTACGTAACCTCAATG 1990
Qy 1801 GTCTGGGATCGGGAAGTGGCGTGTCAAGTGTGGAATAGCGGTGGCGCTGTATCCCGG 1860
Db 1991 GTCTGGGATCGGGAAGTGGCGTGTCAAGTGTGGAATAGCGGTGGCGCTGTATCCCGG 2050
Qy 1861 TGGGAGTACTATGACTCAGGTATACAGAACGTTATAGTGGTCTCCCAACTCCAGAGAC 1920
Db 2051 TGGGAGTACTATGACTCAGGTATACAGAACGTTATAGTGGTCTCCCAACTCCAGAGAC 2110
Qy 1921 AACCTTGACCATTTACAGAAATTCACAGTCATGAGCAGAGCTGAAAATTTTAAACAAGTT 1980
Db 2111 AACCTTGACCATTTACAGAAATTCACAGTCATGAGCAGAGCTGAAAATTTTAAACAAGTT 2170
Qy 1981 GAGTACTCTCTTATTATTCAGTAAACAGCAGATGATTAACGTTTCACTTTTCAGCAGTCAGCTCAG 2040
Db 2171 GAGTACTCTCTTATTATTCAGTAAACAGCAGATGATTAACGTTTCACTTTTCAGCAGTCAGCTCAG 2230
Qy 2041 ATCTCCAAAGCCCTGGTGTGATGTTGGAGTGGATTTTCAGGCAATGTGGTATATGATGAA 2100
Db 2231 ATCTCCAAAGCCCTGGTGTGATGTTGGAGTGGATTTTCAGGCAATGTGGTATATGATGAA 2290
Qy 2101 GACCATGGAATAGCTAGCAGCAGCAGACCAACATATATATACCAATGAGCCACTTC 2160
Db 2291 GACCATGGAATAGCTAGCAGCAGCAGACCAACATATATATACCAATGAGCCACTTC 2350
Qy 2161 ATAAAAAATGTTTCTCTTTACCT 2184
Db 2351 ATAAAAAATGTTTCTCTTTACCT 2374

RESULT 5

US-09-023-655-1026
; Sequence 1026, Application US/09023655
; Patent No. 6607879

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1026:
SEQUENCE CHARACTERISTICS:
LENGTH: 2924 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g180082

US-09-023-655-1026

Query Match 99.9%; Score 2180.8; DB 3; Length 2924;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2182; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGTCGCAAACTTACACTTAACCTGATTAATAAATACTTATAGACTGAAGTTATAC 60
Db 125 AGTCGCAAACTTACACTTAACCTGATTAATAAATACTTATAGACTGAAGTTATAC 184
Qy 61 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAGAAAATAATCTTGGTA 120
Db 185 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAGAAAATAATCTTGGTA 244
Qy 121 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
Db 245 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 304
Qy 181 GGACATTTCTCAATGATTTTCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC 240
Db 305 GGACATTTCTCAATGATTTTCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC 364
Qy 241 AACTACGTGAAGCAATGGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300

Db 365 AACTAGGTGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 424
Qy 301 AAAAGCGAGCTGATTTACAGAGAGAGAGATTCCAAACAACACACACAGTGGGTCACTGGTCA 360
Db 425 AAAAGCGAGCTGATTTACAGAGAGAGAGATTCCAAACAACACACACAGTGGGTCACTGGTCA 484
Qy 361 CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACATGACATTTATGTTAAATTTGAACCA 420
Db 485 CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACATGACATTTATGTTAAATTTGAACCA 544
Qy 421 AATTTACCAAGTTACAGAAATCACATGGAGCGGGAAGAGATATATAATATAATGGAATA 480
Db 545 AATTTACCAAGTTACAGAAATCACATGGAGCGGGAAGAGATATATAATATAATGGAATA 604
Qy 481 ACTGACTGGGTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGCTGCTGGTGGTCTCCA 540
Db 605 ACTGACTGGGTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGCTGCTGGTGGTCTCCA 664
Qy 541 AACGGCACTTTTATGACATATGCCCAATTTAAACGACACAGAGTCCCACTTATTAATATAC 600
Db 665 AACGGCACTTTTATGACATATGCCCAATTTAAACGACACAGAGTCCCACTTATTAATATAC 724
Qy 601 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 660
Db 725 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 784
Qy 661 GCAGGAGCTGTGAATCCCACTGTAAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCA 720
Db 785 GCAGGAGCTGTGAATCCCACTGTAAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCA 844
Qy 721 GTCACCAATGCAATTCCTCATACAAATCACTGCTCTCTCTCTATGTTGATAGGGATCAC 780
Db 845 GTCACCAATGCAATTCCTCATACAAATCACTGCTCTCTCTCTATGTTGATAGGGATCAC 904
Qy 781 TACTTGTGTGATGACATGGGCAACACAGAAAGATTTCTTTGCACTGCTCAGTGGCTCAGGAG 840
Db 905 TACTTGTGTGATGACATGGGCAACACAGAAAGATTTCTTTGCACTGCTCAGGAGG 964
Qy 841 ATTCAGAACTATTCGGTCACTGGATATTTGTGACTATGATGAATCCAGTGGAGAGTGAAC 900
Db 965 ATTCAGAACTATTCGGTCACTGGATATTTGTGACTATGATGAATCCAGTGGAGAGTGAAC 1024
Qy 901 TGCTTAGTGGACGGCAACACATTTGAATGATGACTACTGCTGGGTTGGAAGATTTAGG 960
Db 1025 TGCTTAGTGGACGGCAACACATTTGAATGATGACTACTGCTGGGTTGGAAGATTTAGG 1084
Qy 961 CCTTCAGAACTATTTACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 1085 CCTTCAGAACTATTTACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1144
Qy 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAGAGTGCACATTTATTACA 1080
Db 1145 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAGAGTGCACATTTATTACA 1204
Qy 1081 AAAGGCACTGGGAAGTCACTGGGATAGAAGCTCTAACCCAGTGATTTATCTATCTACATTT 1140
Db 1205 AAAGGCACTGGGAAGTCACTGGGATAGAAGCTCTAACCCAGTGATTTATCTATCTACATTT 1264
Qy 1141 AGTAATGATATAAAGGATCCAGGAGGAGAGGATCTTTTATAAAATCCAACTTATTGAC 1200
Db 1265 AGTAATGATATAAAGGATCCAGGAGGAGGAGGATCTTTTATAAAATCCAACTTATTGAC 1324
Qy 1201 TATACAAAGTGACATGCTCAGTTGTGAGCTGAAATCCGGAAGAGTGTGAGTACTATTCT 1260
Db 1325 TATACAAAGTGACATGCTCAGTTGTGAGCTGAAATCCGGAAGAGTGTGAGTACTATTCT 1384
Qy 1261 GTGTCACTTACGTAAGAGGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTGGTCTGCC 1320
Db 1385 GTGTCACTTACGTAAGAGGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTGGTCTGCC 1444
Qy 1321 CTCTATCTTACACAGAGGCTGATGATATAAGGCTGAGAGTCTCGGAAGACATTTCA 1380
Db 1445 CTCTATCTTACACAGAGGCTGATGATATAAGGCTGAGAGTCTCGGAAGACATTTCA 1504

Qy 1381 GCTTTGGATAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAAGTGGACTTCATTATT 1440
Db 1505 GCTTTGGATAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAAGTGGACTTCATTATT 1564
Qy 1441 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGATATAATCCAAAG 1500
Db 1565 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGATATAATCCAAAG 1624
Qy 1501 AAATATCTCTTACTTATAGATGTTGATGACGCCCATGTAGTCAAAAAAGACAGACTGTC 1560
Db 1625 AAATATCTCTTACTTATAGATGTTGATGACGCCCATGTAGTCAAAAAAGACAGACTGTC 1684
Qy 1561 TTCAAGACTGAACTGGGCCACTTTACTTTGCAAGCAGACAGAAAAACATTTATAGTAGTACTGCTTT 1620
Db 1685 TTCAAGACTGAACTGGGCCACTTTACTTTGCAAGCAGACAGAAAAACATTTATAGTAGTACTGCTTT 1744
Qy 1621 GATGGCAGAGAGAGTGGTTTACCAAGGAGATGAAGATCATGATGCATCAACAGAAAGACTG 1680
Db 1745 GATGGCAGAGAGAGTGGTTTACCAAGGAGATGAAGATCATGATGCATCAACAGAAAGACTG 1804
Qy 1681 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1740
Db 1805 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1864
Qy 1741 GTGACACAAACCAATTTGCAATTTGGGGCTGGTCAATGAGGAGTACGTAACCTCAATG 1800
Db 1865 GTGACACAAACCAATTTGCAATTTGGGGCTGGTCAATGAGGAGTACGTAACCTCAATG 1924
Qy 1801 GTCCTGGGATCGGGAAGTGGCGTGTTCAGATGAGGAATAGCCGTGGCGCTGTATCCCGG 1860
Db 1925 GTCCTGGGATCGGGAAGTGGCGTGTTCAGATGAGGAATAGCCGTGGCGCTGTATCCCGG 1984
Qy 1861 TGGGAGTACTATGACTCAGTGTACACAGAACTGTTACATGGGTCTCCCAACTCCAGAGAC 1920
Db 1985 TGGGAGTACTATGACTCAGTGTACACAGAACTGTTACATGGGTCTCCCAACTCCAGAGAC 2044
Qy 1921 AACCTTGACCAATTCAGAAATTTCAACAGTCAATGAGCAGAGCTGAAAATTTTAAACAAGTT 1980
Db 2045 AACCTTGACCAATTCAGAAATTTCAACAGTCAATGAGCAGAGCTGAAAATTTTAAACAAGTT 2104
Qy 1981 GAGTACCTCTTATTCATGGAACAGCAGATGAATGATGATGATGATGATGATGATGATGATGAT 2040
Db 2105 GAGTACCTCTTATTCATGGAACAGCAGATGAATGATGATGATGATGATGATGATGATGATGAT 2164
Qy 2041 ATCTCCAAAGCCCTGGTGGATGTTGGAGTGGATTTCCAGGCAATGTTGATGATGATGATGAT 2100
Db 2165 ATCTCCAAAGCCCTGGTGGATGTTGGAGTGGATTTCCAGGCAATGTTGATGATGATGATGAT 2224
Qy 2101 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2160
Db 2225 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2284
Qy 2161 ATAAAAACAATGTTTCTCTTTACCT 2184
Db 2285 ATAAAAACAATGTTTCTCTTTACCT 2308

RESULT 6

PCT-US93-07923-1

; Sequence 1, Application PC/TUS9307923

; GENERAL INFORMATION:

; APPLICANT: Morimoto, Chikao

; APPLICANT: Schloeman, Stuart P.

; APPLICANT: Tanaka, Toshiaki

; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: U.S.A.

```

; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; OPERATING SYSTEM: IBM PS/2 Model 502 or 55SX
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 08-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2924
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US93-07923-1

Query Match 99.9%; Score 2180.8; DB 6; Length 2924;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2182; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAAACTTACACTCTACTGATTACTTAAATAATCTATAGACTGAAGTTATAC 60
DB 125 AGTCGCAAACTTACACTCTACTGATTACTTAAATAATCTATAGACTGAAGTTATAC 184
QY 61 TCCTTAAGATGGATTTGAGATCATGAATATCTCTACAAACAAAGAAATATATCTGGTA 120
DB 185 TCCTTAAGATGGATTTGAGATCATGAATATCTCTACAAACAAAGAAATATATCTGGTA 244
QY 121 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGTTT 180
DB 245 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGTTT 304
QY 181 GGACATTTCTATCAATGATTATTCAATATCTCTGATGGGAGTTTATTTCTTTAGAAATAC 240
DB 305 GGACATTTCTATCAATGATTATTCAATATCTCTGATGGGAGTTTATTTCTTTAGAAATAC 364
QY 241 AACTAGTGAAGCAATGGAGGCAATTCCTACACAGCTTCTATATGACATTTATGATTTAAAT 300
DB 365 AACTAGTGAAGCAATGGAGGCAATTCCTACACAGCTTCTATATGACATTTATGATTTAAAT 424
QY 301 AAAAGGCAGCTGATTACAGAGAGAGGATTCCAAACAAACACAGCTGGGTGACATGGTCA 360
DB 425 AAAAGGCAGCTGATTACAGAGAGAGGATTCCAAACAAACACAGCTGGGTGACATGGTCA 484
QY 361 CCAGTGGGTGCATAAAATGGGCATATGTTTGGAAACATGACATTTATGTTAAATTTGAACCA 420
DB 485 CCAGTGGGTGCATAAAATGGGCATATGTTTGGAAACATGACATTTATGTTAAATTTGAACCA 544
QY 421 AATTTACCAAGTTTACAGATTCATGACGGGAAAGAGATATATATATATATGATTAATGGAATA 480
DB 545 AATTTACCAAGTTTACAGATTCATGACGGGAAAGAGATATATATATATATGATTAATGGAATA 604
QY 481 ACTGACTGGTTTATGAGGAGGAGTCTTTCAGTGGCTACTCTGCTGCTGGTGGTCTCCA 540
DB 605 ACTGACTGGTTTATGAGGAGGAGTCTTTCAGTGGCTACTCTGCTGCTGGTGGTCTCCA 664
QY 541 AACGGCAGCTTTTATGACATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTATGAAATAC 600

```

```

DB 665 AACGGCAGCTTTTATGACATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTATGAAATAC 724
QY 601 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCCATATCCNAAG 660
DB 725 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCCATATCCNAAG 784
QY 661 GCAGGAGCTGTGAATCCCAACTGTAAAGTCTTTGTTGTAATAACAGACTCTCTCAGCTCA 720
DB 785 GCAGGAGCTGTGAATCCCAACTGTAAAGTCTTTGTTGTAATAACAGACTCTCTCAGCTCA 844
QY 721 GTCCACCAATGCAACTTCCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
DB 845 GTCCACCAATGCAACTTCCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 904
QY 781 TACTTGTGTGATGACATGGGCAACACAAAGAAAGAAATTTCTTTTCAGTGGGTCTCAGGAGG 840
DB 905 TACTTGTGTGATGACATGGGCAACACAAAGAAAGAAATTTCTTTTCAGTGGGTCTCAGGAGG 964
QY 841 ATTTCAGAACTATTTCGGTTCATGGATATTGTTGACTATGATGAATCCAGTGGAAAGATGGAAC 900
DB 965 ATTTCAGAACTATTTCGGTTCATGGATATTGTTGACTATGATGAATCCAGTGGAAAGATGGAAC 1024
QY 901 TGTCTTAGTGGCACGGCAACACATTCGAAATGAGTACTACTGCTGGTGGTGGAAAGATTTAGG 960
DB 1025 TGTCTTAGTGGCACGGCAACACATTCGAAATGAGTACTACTGCTGGTGGTGGAAAGATTTAGG 1084
QY 961 CTTTCAGAACTCTTATTTACCTTTGATGGTAAATAGTCTTCTCAAGATCATCAGCAATGAA 1020
DB 1085 CTTTCAGAACTCTTATTTACCTTTGATGGTAAATAGTCTTCTCAAGATCATCAGCAATGAA 1144
QY 1021 GAAGGTTACAGACACATTTGCTATTTCCTCAATATAGATAAAGAAAGCTGCACATTTATTACA 1080
DB 1145 GAAGGTTACAGACACATTTGCTATTTCCTCAATATAGATAAAGAAAGCTGCACATTTATTACA 1204
QY 1081 AAAGSCACTCGGAGTCACTCGGATAGAAGCTCTAACAGTGAATTAATCTATCTATCTACATT 1140
DB 1205 AAAGSCACTCGGAGTCACTCGGATAGAAGCTCTAACAGTGAATTAATCTATCTATCTACATT 1264
QY 1141 AGTAATGAATATAAGGAATGCCAGGAGGAGGAGTCTTTATAAATCCAACTTATTGAC 1200
DB 1265 AGTAATGAATATAAGGAATGCCAGGAGGAGGAGTCTTTATAAATCCAACTTATTGAC 1324
QY 1201 TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGGTTGTCAGTACTATTCT 1260
DB 1325 TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGGTTGTCAGTACTATTCT 1384
QY 1261 GTGTCATTCAGTAAAGGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCGGCTGCCCC 1320
DB 1385 GTGTCATTCAGTAAAGGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCGGCTGCCCC 1444
QY 1321 CTCCTATCTTACACAGCAGCGTGAATGATATAAGGCTGAGAGTCTCTGGAAGACAAATTTCA 1380
DB 1445 CTCCTATCTTACACAGCAGCGTGAATGATATAAGGCTGAGAGTCTCTGGAAGACAAATTTCA 1504
QY 1381 GCTTTGGGATAAAATGCTGCAGAACTCCAGATGCCCTCCAAATAAACTGGACTTCAATATT 1440
DB 1505 GCTTTGGGATAAAATGCTGCAGAACTCCAGATGCCCTCCAAATAAACTGGACTTCAATATT 1564
QY 1441 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTTGGCTCTCTCAATTTTGATAAATCCAAAG 1500
DB 1565 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTTGGCTCTCTCAATTTTGATAAATCCAAAG 1624
QY 1501 AAATATCTCTACTATTAGATGTGTATGCGGCCCATGTGTAGTCAAAAGCAGACACTGTCTC 1560
DB 1625 AAATATCTCTACTATTAGATGTGTATGCGGCCCATGTGTAGTCAAAAGCAGACACTGTCTC 1684
QY 1561 TTTCAGACTGAACTGGGCCACTTTACCTTGCAGACACAGAAAAATTTATATAGTACTAGCTTTT 1620
DB 1685 TTTCAGACTGAACTGGGCCACTTTACCTTGCAGACACAGAAAAATTTATATAGTACTAGCTTTT 1744
QY 1621 GATGCGAGAGGAGTGGTTTACCAAGGAGATTAAGATCATGATGCAATCAACAGAGACTGTG 1680
DB 1745 GATGCGAGAGGAGTGGTTTACCAAGGAGATTAAGATCATGATGCAATCAACAGAGACTGTG 1804

```

```

QY 1691 GGAACATTGAGTTGAAGATCAAAATTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 1740
DB 1805 GGAACATTGAGTTGAAGATCAAAATTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 1864
QY 1741 GTGGACAAACAAAGATTTGCAATTTGGGGCTGGTCAATATGAGGGGTACGTAACCTCAATG 1800
DB 1865 GTGGACAAACAAAGATTTGCAATTTGGGGCTGGTCAATATGAGGGGTACGTAACCTCAATG 1924
QY 1801 GTCTGGGATCGGGAAGTGGCGTGTTCAGATGTGGAATAGCCGTGGCGCTGTATCCCGG 1860
DB 1925 GTCTGGGATCAGGAAGTGGCGTGTTCAGATGTGGAATAGCCGTGGCGCTGTATCCCGG 1984
QY 1861 TGGGAGTACTATGACTCAGTGTGACACAGAACTTATACATGGGTCTCCCACTCCAGAAC 1920
DB 1985 TGGGAGTACTATGACTCAGTGTGACACAGAACTTATACATGGGTCTCCCACTCCAGAAC 2044
QY 1921 AACCTTGACCAATACAGAAATTTCAACAGTCAATGAGCAGAGCTGAAATTTTAAACAAGTT 1980
DB 2045 AACCTTGACCAATACAGAAATTTCAACAGTCAATGAGCAGAGCTGAAATTTTAAACAAGTT 2104
QY 1981 GAGTACCTCTCTTATTCATGGAACAGCAGATGATACGTTTCACTTTTCACTGCTCAGCTCAG 2040
DB 2105 GAGTACCTCTCTTATTCATGGAACAGCAGATGATACGTTTCACTTTTCACTGCTCAGCTCAG 2164
QY 2041 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATACCTGATGAA 2100
DB 2165 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATACCTGATGAA 2224
QY 2101 GACATGGAATAGCTAGCAGACAGCAGACACCAACATATATATATATATATATATATATAT 2160
DB 2225 GACATGGAATAGCTAGCAGACAGCAGACACCAACATATATATATATATATATATATATAT 2284
QY 2161 ATAAACAATCTTCTCTTTTACCT 2184
DB 2285 ATAAACAATCTTCTCTTTTACCT 2308

```

RESULT 7

US-08-230-491A-1

; Sequence 1, Application US/08230491A

; Patent No. 5587299

; GENERAL INFORMATION:

; APPLICANT: Rettig, Wolfgang J.; Scanlan, Matthew J.;

; APPLICANT: Garin-Chesa, Pilar; Old, Lloyd J.

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR

; TITLE OF INVENTION: FIBROBLAST ACTIVATION PROTEIN AND USES

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FELPE & LYNCH

; STREET: 805 THIRD AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE 3.5 inch 1.2 MB STORAGE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: WORDPERFECT - ASC II

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/230,491A

; FILING DATE: 20-APRIL-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 5587299man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 330

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2815 Base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; US-08-230-491A-1

Query Match 34.4%; Score 751; DB 2; Length 2815;

Best Local Similarity 60.9%; Pred. No. 1e-218;

Matches 1320; Conservative 0; Mismatches 825; Indels 24; Gaps 5;

```

QY 15 CACTCTAATGATTTACTTAAATAATCTTATAGACTGAAGTTATCTCTTAAGATGAT 74
DB 334 CACACTGAAGGATATTTTAAATGGAACAATTTTCTTATAAAACAATTTTTCAAAACCTGGAT 393
QY 75 TTCAGATCATCAATATCTCTACAAACAAGAAATAATCTTTGGTATTCAATGCTGAATA 134
DB 394 TTCAGGACAGAAATATCTTCATCAATCTGAGAGTAAACAATATAGTACTTTTAAATATTGA 453
QY 135 TGGAAACAGCTCAGTTTTCTTTGGAGAAACAGTACATTTTGATGAGTTTGGACATCTTATCAA 194
DB 454 AACAGGACAATCATATACCAATTTTGGAGTAAATAGAACCATGAAAGTGTGAATGCTTCAA- 512
QY 195 TGATTTATCAATATCTCTGATGGGCAAGTTTATTTCTCTTAGAATACAACTAGGTGAAGCA 254
DB 513 --ATTACGGCTTATCAGCTGATCGGCAATTTGTATATCTAGAAAAGTGAATTTTCAAAGCT 570
QY 255 ATGAGGACATTTCTTACACAGCTTCTATGACATTTATGATTTTAAATAAAGCGAGCTGAT 314
DB 571 TTGAGATATCTTTTACACAGCAATATATCATCTATGACCTTAGCAATGGAGAAATTTGT 630
QY 315 TACAGAAGAGAGGATTTCCAAACAACACACAGTGGGTCACTGGTCCAGGTGGGTCTATAA 374
DB 631 AAGAGGAATGAGCTTCTCTGCTCCAAATTCAGTATTTATGCTGGTGGCTGTGGGAGTAA 690
QY 375 ATTGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACCAAAATTTTACCAAGTTA 434
DB 691 ATTAGCATATGCTCTATCAAAACAATATCTATTGGAACAAGACACGAGAGATCCACCTTT 750
QY 435 CAGAAATCAGATGGAAGGGAAGAGATATAATATAATGGAATTAACCTGAGTGGTTTAA 494
DB 751 TCAAAATACATTTAATGGAAGAGAAATAAATAATTTAATGGAATCCCAAGCTGGGTTTAA 810
QY 495 TGAAGAGAGAGTCTTTCAAGTGCCTACTCTGCTCTGTGGTGGTCTCCAAACGGCAGCTTTTT 554
DB 811 TGAAGAGGAATGCTTCTTCAAAATATGCTCTCTGCTGGTGGTCTCTTAATGGAATTTTT 870
QY 555 AGCATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTGAATACTCTCTTCTACTCTGA 614
DB 871 GGCATATGCGGAATTTAATGATAAGGATATACCAAGTATTGCTCTATTCTTATTATGGCGA 930
QY 615 TGAGTCACTGCAGTACCCAAAGACTGTACGGGTTTCCATATCCAAAGGCGAGGCTGTGAA 674
DB 931 TGA-----ACAATATCTTAGAACAAATAAATTTTCCATACCCCAAGGCTGGAGCTTAAGAA 984
QY 675 TCCAACTGTAAGTCTTTTGTGTAATAACAGACTCTCTCAGCTCAGTCAACAATGCAAC 734
DB 985 TCCCGTGTTCGGATATTTATTATCGATAC-----CACTTACCTCGGTATGTAGG 1035
QY 735 TTCCATACAAATCACTGCTCTGCTTCTTATGTTGATGGGATCACTACTTGTGTGATGT 794
DB 1036 TCCCCAGGAAGTGCCTGTTCCAGCAATGATAGCCCTCAAGTGAATTTATTTTCAAGTTGGCT 1095
QY 795 GACATGGGCAACACAGAAAGAAATTTCTTTGCACTGAGTCTAGGAGGATTCAGAACTATTC 854
DB 1096 CAGTGGGTACTGATGAACAGAGTATGTTTGCAGTGGCTAAAAGAGTCCAGAAATGTTTC 1155
QY 855 GGTCAATGATATTTGTGACTATGATGAATCCAGTGGAGAGTGAAGTCTTAGTGGCAGC 914
DB 1156 GGTCTGCTCTATATGACTTTCAGGGAAGACTGCGACACATGGGATTTGTCCAAAGACCA 1215
QY 915 GCACACATTTGAATGATGATCTACTGCTGGGTTTGGAGAGATTTTAGGCTTCCAGAACCTCA 974

```


Qy	1272	TAAGAGCGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCGGTCTGCCCTCTATATCTCT	1331
Db	1576	CGACTACGCCAAGTACTATGCATTTGCTCTACGGCCAGGCAATCCCCATTTCCACCCCT	1635
Qy	1332	ACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCCTGGAAGACAAATTCAGCTTTGGATAA	1391
Db	1636	TCATGATGGACGCACTGATCAGAAATTAATAATCCTGGAAGAAAACAAGGAATTTGGAAAA	1695
Qy	1392	AATGCTCGAGAAATGTCAGATGCGCTCCAAAAAATCGGACTTCATTATTTTGAATGAAC	1451
Db	1696	TGCTTTGAAAAATATCCAGCTGCTTAAGAGGAAAAATTAAGAAACTTGAAGTAGATGAAT	1755
Qy	1452	AAAAATTTTGTGTACAGATGATCTTGCTCTCTCATTTTGTATAAATCCNAGAAATATCTCT	1511
Db	1756	TACTTTATGGTACAAAGATGATTTCTTCTCTCAATTTTGACAGATCAAGAAGATATCCCTT	1815
Qy	1512	ACTATTAGATGTGTATCAGCGCCCATGTAGTCAAAAGACAGACACTGTCTTCAGACTGAA	1571
Db	1816	GCYAAATTCAGTGTATGGTGGTCCCTGCACTCAGAGTGTAAAGTCTGTATTTGCTGTAA	1875
Qy	1572	CTGGGCCACTTACCTTTGCAAGCAACAGAAAAATTTATAGTAGTACGTCTTGTATGGCAGAGG	1631
Db	1876	TTGGATATCTTATCTTCCAGTAGGAAGGATGGTCAATTCCTTGGTGCATGGTCGAGG	1935
Qy	1632	AAGTGGTTACCAAGGAGATGAAGTATCATGCAATCAACAGAGACTCGGAACATTTGA	1691
Db	1936	AACAGCTTTTCCAAAGGTGACAAATCCCTCTATAGCAGTGTATCGAAAGCTGGGTGTATGA	1995
Qy	1692	AGTTGAAGATCAAAATGAAGCAGCCAGACAAATTTTCAAAAATCGGATTTCTGTGACAAACAA	1751
Db	1996	AGTTGAAGACCAAGTTACAGCTGTACAGAAATTTCAAGAAATGGGTTTCATTGATGAAA	2055
Qy	1752	ACGAATTTGCAATTTGGGGCTGGTCATATGAGAGGTACGTAACTCAATGGTCTCTGGGATC	1811
Db	2056	AAGAAATAGCCATATGGGGCTGGTCTCATGAGGATACGTTTCATCATCTGGCCCTTGCATC	2115
Qy	1812	GGGAGTGGCGTGTTCAGTGTCTGAATAGCCGTGGCCCTGTATCCGGTGGGAGTACTA	1871
Db	2116	TGGAACTGGTCTTTTCAAAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAAATATTA	2175
Qy	1872	TGACTCAGTGTACACAGAACGTTTACATGGGTCTCCAACTCCAGAGACAACTTCACCA	1931
Db	2176	CGCGTCTGTCTACACAGAGATTTCAATGGGTCTCCCAACAAAGAGATGATTAATCTTGAGCA	2235
Qy	1932	TTACAGAAATTCACACAGTCTATGACGAGAGCTGAAAAATTTTAAACAAGTTGAGTACCTCT	1991
Db	2236	CTATAAGAAATTCACATGTGTATGCAAGAGAGCAGAAATATTTTCAGAAAAATGTAGACTATCTCT	2295
Qy	1992	TATTTATGAGAACAGCAGATGATTAACGTTTCACTTTTACGAGCTCAGCTCAGATCTCCAAAGC	2051
Db	2296	CATCCACGGAAACAGCAGATGATTAATGTGCATTTTCAAAACTTCAGCACAGATTTGCTAAAGC	2355
Qy	2052	CCTGGTTCGATGTTGGAGTGGATTTCCAGGCCAAATGTGTGTATCTACTGATGAAGACCATGGAT	2111
Db	2356	TCTTGGTTAATGCAACAGTGGATTTCCAGGCATGTGGTACTCTTGACCAAGAACCAACGGCTT	2415
Qy	2112	AGCTAGCAGCACAGCACCAACAT	2171
Db	2416	ATCCGGCTGTCCACGAAC---CACTTATACACCCATGATGACCCACTTCTCTAAAGCAGTG	2472
Qy	2172	TTTCTCTTTT	2180
Db	2473	TTTCTCTTTT	2481

RESULT 9
US-08-940-391-1
; Sequence 1, Application US/08940391
; Patent No. 5965373
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION

TITLE OF INVENTION: PROTEIN ALPHA, AND USES THEREOF

; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Felfe & Lynch
 ; STREET: 805 Third Avenue
 ; CITY: New York City
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10022
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/940,391
 ; FILING DATE: 01-OCT-1997
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/619,280
 ; FILING DATE: 18-MARCH-1996
 ; APPLICATION NUMBER: 08/230,491
 ; FILING DATE: 20-APRIL-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hanson, No. 5965373man D.
 ; REGISTRATION NUMBER: 30,946
 ; REFERENCE/DOCKET NUMBER: LUD 5330.1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 688-9200
 ; TELEFAX: (212) 838-3884
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2815 Base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; US-08-940-391-1

Query Match 34.4%; Score 751; DB 2; Length 2815;
 Best Local Similarity 60.9%; Pred. No. 1e-218;
 Matches 1320; Conservative 0; Mismatches 825; Indels 24; Gaps 5;

QY	15	CACCTCTAAGTCTTCTTAAATACTTATGAGCTGAAAGTTATATCTCTCTTAAGTGGAT	74
DB	334	CACACTGAAGGATATTTTAAATGGAACATTTCTTATAAACAATTTTTCGAACTGGAT	393
QY	75	TTTCAGATCATGAATATCTCTACAAAACAGAAAATAATATCTTGGTATTCGAATGCTGAATA	134
DB	394	TTTCAGGACAAGAATATCTTCATCAATCTGCAGATACAAATATAGTACTTTTATAATATTGA	453
QY	135	TGMAAACAGCTCGTTTCTTGGAGAACAGTACATTTTGATGAGTTTGGACATTTCTATCAA	194
DB	454	AAACAGGACAATCATATACCAATTTTGAGTAATAGAACCATGAAAAGTGTGAATGCTTCAA	512
QY	195	TGATTTATTCATATCTCTGATGGGAGTTTATCTCTTAGAATACAACTACGTGAAGCA	254
DB	513	--ATTACGGCTTATCACTGATCGGCAATTTGATATCTAGAAAGTATTTTCAAAGCT	570
QY	255	ATGGAGGCATTCCTACACAGCTTTCATATGACATTTTATGATTTTAAATATAAAGGCAGCTGAT	314
DB	571	TTTGAGATATCTTTACACAGCAACATATATACATCTATGACCTTTAGCAATGGAGAAATTTGT	630
QY	315	TACAGAAGAGAGATTTCCAAAACACACAGTGGGTCACATGGTCAACAGTGGGTGATTA	374
DB	631	AAGAGGAAATGAGCTTCTCTGCTCAAATTCAGTATTTTATGCTGGTCCCTGTTGGGAGTAA	690
QY	375	ATTGGCATATGTTTGGAAACAATGACATTTTATGTTTAAATTTGAACCAAAATTTACCAAGTTA	434
DB	691	ATTAGCATATGCTATCAAAAACATATCTATTTTGAACCAAGACAGGAGATCCACCTTT	750
QY	435	CAGAATCACATGGACGGGAAAGAGATATAATATATAATATGAATTAATGGAATTAATGCTGGTTTA	494
DB	751	TCAAAATAACATTTAATGGAAGAGAAAATAAATAATTTAATGGAATCCAGACTGGGTTTA	810

QY	495	TGAAGAGGAAGTCTTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCAAAGCGCACTTTTT	554
DB	811	TGAAGAGGAATGCTTCTTACAAAATATGCTCTCTGTGGTGGTCTCTTAATGGAATAATTT	870
QY	555	AGCATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTTGAATACTCTCTTACTCTGA	614
DB	871	GGCATATGGGAATTTAATGATAAGGATATACCACTTATTCCTATTTCTTATATGGCGA	930
QY	615	TGAGTCACTGCAGTACCCCAAGACTGTACGGGTTCCATATCCAAAGGCGAGGCTGTGAA	674
DB	931	TGA-----ACAATATCTTAGAACAAATAATTCATACCCAAAGGCTGGAGCTAAGAA	984
QY	675	TCCAACTGTAAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCAGTCAACCAATCAAC	734
DB	985	TCCCGTTGTTCCGATATTTATTCGTATC-----CACTTACCCTGCGTATGTAGG	1035
QY	735	TTCCATACAAAATCACTGCTCTCTCTTCTATGTTGATAGGGGATCACTACTGTGTGATGT	794
DB	1036	TCCCAGGAAGTGCCTGTTCCAGCAATGATAGCTCAAGTGATTTATTTTTCAGTTGGCT	1095
QY	795	GACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGCTCAGGAGGATTCAGAACTATTC	854
DB	1096	CACGTGGGTTACTGATGAACGAGTATGTTTGCAGTGGCTTAAAGAGAGTCCAGAACTTTC	1155
QY	855	GGTCATGGATATTTGCTGACTATGATGAATCCAGTGGAGATGGAATGCTTATGTGCAAG	914
DB	1156	GGTCTCTGTATATGCTGCTTCCAGGAGAGACTGGCAGACATGGATTTGCCAAGACCCA	1215
QY	915	GCAACACATTTGAAATGAGTACTTACTGGCTGGGTTGGAGATTTAGGCCCTTCAGAACCTCA	974
DB	1216	GGAGCATATAGAAGAAAGCAGAACTGGATGGCTGGTGGATTTCTTTGTTTCAAGACCAGT	1275
QY	975	TTTTACCTTGTGATGTAATAGCTTCTACAAGATCATCAGCAATGAAGAGTTTACAGACA	1034
DB	1276	TTTCAGCTATGATGCCATTTTCGTACTACAAAATATTTAGTCAAGAGGATGGCTTACAAA	1335
QY	1035	CATTTGCTATTTTCCAAATAGATAAAAAGACTGACATTTATTAACAAAAGGACCTGGGA	1094
DB	1336	TATTCATATATCAAAAGACACTGTGGAAAATGCTATTCAATTTACAAAGTGGCAAGTGGG	1395
QY	1095	AGTCATCGGGATAGAAGCTCTAACCAAGTGAATTTATCTACTACATTTAGTAAATATAA	1154
DB	1396	GGCCATAAATATATTTTCAGAGTAAACACAGGATTCCTGTTTATTTCTAGCAATGAATTTGA	1455
QY	1155	AGGAATGCCAGAGAGAGAAATCTTTATAAATCCAACTTATTTGACTA---TACAAAAGT	1211
DB	1456	AGAAATACCTTGGAAAGAAACATCTACAGAAATTAGCATTTGGAAGCTATCTCCCAAGCAA	1515
QY	1212	GACATGCTCTCAGTCTGAGCTGAATCCGGAAGGTGTGAGTACTATTCTGTGTCTATTCAG	1271
DB	1516	GAAGTGTGTTACTTCCCATCTAAGGAAGAAAGGTGCCAATATTACAGACGAAGTTTCAG	1575
QY	1272	TAAAGAGCGGAAGTATTTATCAGCTGAGATGTTCCGGTCTGTGCTGCCCCCTCTATACTCT	1331
DB	1576	CGACTACGCCAAGTACTATGCACTTGTCTGCTACGSCCCAGGCATCCCCATTTCCACCCT	1635
QY	1332	ACACAGCAGCGTAATGATATAAGGGCTGAGAGTCTCTGGAGACAAATTCAGCTTTGGATAA	1391
DB	1636	TCATGATGGACGCACATGATCAAGAAATTTAAAAATCTCTGGAAAGAAACAGGAAATTTGAAA	1695
QY	1392	AATGCTGCAAGATGTCAGATGCCCTCCAAAAAATCTGGAATTTTATTTTGAATGAAAC	1451
DB	1696	TGCTTTGAAAATATCCAGCTGCCCTAAGAGGAAATTAAGAAACTTGAAGTAGATGAAAT	1755
QY	1452	AAAAATTTGGTATCAGATGATCTTCCCTCTCATTTTGTGATAAATCCAGAAATATCCTCT	1511
DB	1756	TACTTTATGGTACAAGATGATTTCTTCTCTCAATTTGACAGATCAAGAAGATATCCCTT	1815
QY	1512	ACTATTAGTGTATGACAGGCCCATGTAGTCAAAAAGCAGACACACTGCTTCCAGACTGA	1571
DB	1816	GCTAATTCAGTGTATGGTGGTCCCTGCAGTCAGAGTGTAAAGTCTGTATTTTCTGTTAA	1875

QY 1572 CTGGCCACTTACCTTGCAGCAGCAAGAAACATTAATAGTACCTTGTGATGGCAGAGG 1631
DB 1876 TTGGATATCTTATCTTGAAGTAAGGAAGGATGCTCATTCCTTGGTGGATGGTCGAGG 1935
QY 1632 AAGTGGTTACCAAGGAGATAGATCATGCTGATCAATCAACAGAGCTGGGACATTTGA 1691
DB 1936 AACAGCTTTTCAAGGTGCAAACTCTTATGCACTGATCGAAAGCTGGGTGTTTATGA 1995
QY 1692 AGTTGAAGATCAAAATTTGAAGCAGCCAGCAAAATTTTCAAAATTTGGATTTGGCAACA 1751
DB 1996 AGTTGAAGCAGATTAACGCTGTGCAAAATTTTCAAAATTTGGATTTGGCAACA 2055
QY 1752 ACGAATTCGAATTTGGGCTGTGATATGAGGGGTACGTAACCTCAATGGTCTGGGATC 1811
DB 2056 AAGAATAGCCATATGGGCTGTGCTTATGAGGATACGTTTTCATCACTGGCCCTTGCATC 2115
QY 1812 GGGAGTGGCTGTTCAGTGTGGATAGCCGTGGGCTGTATCCCGTGGGATCTA 1871
DB 2116 TGGAACTGGTCTTTTCAAAATTTGGATATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTA 2175
QY 1872 TGACTCAGTGTACAGCAAGCTTACATGGGTCTCCAACTCCAGAGCAACCTTGACCA 1931
DB 2176 CGGCTCTGTACAGAGAGATTCATGGGTCTCCAAAGAGATATCTTGACCA 2235
QY 1932 TTACAGAAATCAACAGATCATGAGCAGAGCTGAAATTTTAAACAAAGTTGAGTACCTCT 1991
DB 2236 CTATAAGAAATCAACTGTGTGAGCAAGAGCAGATATTTTCAAAATTTGAGATCTCTCT 2295
QY 1992 TATTCATGCAAGCAGATGATAGCTTACCTTTCAGCAGTCTGAGTCTCCAAAGC 2051
DB 2296 CATCCAGCAAGCAGATGATATATGTCACCTTTTCAAAATTTGAGATCTCTCTCT 2355
QY 2052 CTGTGTGATGTGGAGTGGATTTCCAGGCAATGTGTATCTGATGAAGACCATGGAAT 2111
DB 2356 TCTGTTAATGCAAGTGGATTTCCAGGCAATGTGTATCTGATGAAGACCATGGAAT 2415
QY 2112 AGTAGCAGCAGCAGCAACCAATATATATACCCAGATGAGCCACTTTCATAAACATG 2171
DB 2416 ATCCGGCTGTCCAGAAC---CACTTATACCCAGATGAGCCACTTTCATAAAGCAGTG 2472
QY 2172 TTTCTCTTT 2180
DB 2473 TTTCTCTTT 2481

RESULT 10
US-09-265-606-1
; Sequence 1, Application US/09265606
; Patent No. 6846910
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John B.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felife & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280

; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6846910man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2815 Base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-265-606-1
Query Match 34.4%; Score 751; DB 3; Length 2815;
Best Local Similarity 60.9%; Pred. No. 1e-218;
Matches 1320; Conservative 0; Mismatches 825; Indels 24; Gaps 5;
QY 15 CACTCTAATGATTAATTAATAAATCTTATAGACTGAGTTATCTCTTAAGATGAT 74
DB 334 CACACTGAAGGATATTTTAAATGGAACATTTTCTTATAAAACATTTTTCCTCAAACTGGAT 393
QY 75 TTCAGATCATCAATATCTCTACAAACAGAAATAAATATCTTGGTATTCAATGCTGAATA 134
DB 394 TTCAGGACAGATATCTTCAATCTGAGATTAACATATAGTACTTTTATATATTGA 453
QY 135 TGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTTGATGAGTTTGGACATCTTATCAA 194
DB 454 AACAGGCAATCATATACCATTTTGGAGTAATAGAACATGAAGTGTGAATGCTTCAA- 512
QY 195 TGATTTATCAATATCTCTGATGGCAGTTTATTTCTCTTGAATACAACTAGTGAAGCA 254
DB 513 --ATTACGGCTTATCAGCTGATCGGCAATTTGTATATCTAGAAAGTGTATTTCAAAGCT 570
QY 255 ATGAGGACATTCCTACACAGCTTCATATGACATTTATGATTTAAATAAAGCAGCTGAT 314
DB 571 TTGAGGATATCTTACAGCAACATATATATCTATGACCTTAGCAATGAGAAATTTGT 630
QY 315 TACAGAAAGAGAGATTCCAAACAAACACACAGTGGGTCACTGGTCCACAGTGGGTCAATA 374
DB 631 AAGAGGAAATGAGCTTCTCTGCTCAATTCAGTATTTATGCTGTGCTGCTGTGGAGTAA 690
QY 375 ATTGGCATATGTTTGGAAACATGATGATTTATGTTAAATTTGAACCAAAATTTACCAAGTGA 434
DB 691 ATTAGCATATGCTATCAAAACAAATATCTATTTCGAAACAAAGACAGGAGATCCACCTTT 750
QY 435 CAGAAATCAGATGGAGCGGGAAGAGATATAATATAATATAATGAAATTAATGAAATCCAGACTGGGTTTA 494
DB 751 TCAATTAACATTTAATGGAAGAGAAATAAATAATTTAATGAAATCCAGACTGGGTTTA 810
QY 495 TGAAGAGGAAATCTTTCAGTGCCTACTCTGCTCTGCTGGTGGTCTCCAAACCGCACCTTTTTT 554
DB 811 TGAAGAGGAAATGCTTCTCACAAATATGCTCTCTGCTGGTGTCTCTTAATGGAATAATTTT 870
QY 555 AGCATATGCCCAATTTAAACAGACAGAGTCCCACTTATTGAATATCTCTTCTACTCTGA 614
DB 871 GGCATATGCCGAATTTAATGATAAGGATATACCAAGTTATTTCCTTATTATGCGGA 930
QY 615 TGAGTCACTGCAGTACCCCAAGACTGTACGGGTTCCATATCCAAAGGAGGAGCTGTGAA 674
DB 931 TGA-----ACAATATCTAGAACATAAATAATTTCCATACCCCAAGGCTGGAGCTTAAGAA 984
QY 675 TCCAACTGTAAAGTTCTTTGTTGTTAAATACAGACTCTCTCAGCTCAGTCACTCAATATGCAAC 734
DB 985 TCCCGTTGTTCCGATATTTTATTCGATAC-----CACTTACCTGCTGATATGATG 1035
QY 735 TTCCATACAAATCACTGCTCTGCTGCTTCTTATGTTGATAGGGGATCACTACTTGTGTGATGT 794
DB 1036 TCCCGGAGAGTGCCTGTTCAGCAATGATAGCCTCAAGTGTGATTTATTTTTCAGTTGGCT 1095


```
Oy 315 TACAGAGAGAGGATTCCAAACAAACACACAGTGGGTCCATGGTCCACAGTGGGTCTATAA 374
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 631 AAGAGGAAATGAGCTTCTCGTCCAAATTCAGTATTATGCTGGTGGTGGAGTAA 690
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 375 ATTGGCATATGTTGGAAACATGACATTTATGTTAAATTTGAACCAAAATTTACCAAGTTA 434
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 691 ATTAGCATATGCTATCAAAACAAATATCTATTGTAACAAAGACCAAGGAGATCCACCTTT 750
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 435 CAGAAATCAATGCGGGAAGAGATATAATATATAATGGAATAACCTGACTGGTTTA 494
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 751 TCNAATACATTAATGGAAGAAATTAATAATATTAATGGAATCCACAGCTGGTTTA 810
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 495 TGAAGAGAGGCTTTCACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 554
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 811 TGAAGAGGAAATGCTTGCTACAAAATAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 555 AGCATATGCCAAATTTACGACACAGAGTCCACTATTGAATACCTCTCTCTCTCTCTCTCT 614
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 871 GGCATATCGGAAATTTAATGATACGGATATACCAATATTCAGTTATTCCTATTATGGCGA 930
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 615 TGAGTCACTGAGTACCAAGACTGTACGGGTTCCATATCCAAAGCAGGAGCTGTGAA 674
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 931 TGAA-----CAATATCTAGAACAAATAATTTCCATACCAGGCTGGAGCTAAGAA 984
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 675 TCCAACTGTAAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCAGTCACCAATGCCAAC 734
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 985 TCCCGTTGTCGGATATTTATATCGATAC-----CACTTACCCTGGGTATGAG 1035
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 735 TTCATACAAATCACTGCTCTGCTCTATGTTGTAAGGGAATCACTATTGTTGATGTT 794
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1036 TCCCGAGGAGTGGCTGTTCCAGCAATGATAGCCTCAAGCTGATTTATTTTCACTGGCT 1095
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 795 GACATGGCAACACAGAAAGAAATTTCTTCAGTGGCTCAGGAGGATTCAGAACTATTC 854
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1096 CACGTGGTACTGATGAACAGATGTTTCAGTGGCTTAAAGAGGTCAGAAATGTTTC 1155
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 855 GGTCAATGGATATTTGCACTATGATGAATCCAGTGGAGATGGAATCTTTAGTGGCAGC 914
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1156 GGTCTGTCTATATGTGATTCAGGGAAGACTGGCAGACATGGGATTTGCCAAGACCCA 1215
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 915 GCAACATTTGAATGATGACTACTGGCTGGGTTGGAAGATTTAGGCCCTCAGAACTCA 974
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1216 GGAGCATATAGAAGAAAGCAAACTGGATGGGCTGGTGGATCTTTGTTTCAACACCACT 1275
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 975 TTTTACCTTGATGTAATCTTCTCAAGATCATCAGCAATGAGAGGTTTACAGACA 1034
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1276 TTTCAGCTATGCTATTCGTACTACAAATATTTAGTGCAAGGATGGCTACAAACA 1335
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1035 CATTTGCTATTTCAAATAGATAAAAGACTGCACATTTTATTAACAAAGGCACCTGGGA 1094
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1336 TATTCATATATCAAGACACTGTGGAAATGCTATTCAAAATACAGTGGCAAGTGGGA 1395
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1095 AGTCATGGGATAGAGCTTAAACAGTGAATTTCTATCTATCTATCTATCTATCTATCTAT 1154
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1396 GGCCATAAATATTTACAGATTAACACAGGATTCAGTGTGTTTATTTCTAGCAATGAATTGA 1455
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1155 AGGAATCCAGGAGGAGGATCTTTATAAAATCCAACTTATGACTATA---CAAAAGT 1211
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1456 AGAATACCTCGGAAGAGAAACATCTACAGAAATTAGCAATGGAAAGCTATCTCCAAAGAA 1515
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1212 GACATGCTCAGTTGTGAGCTGAAATCCGGAAGGTGTGAGTACTTATCTGTGTCATTCAG 1271
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1516 GAAGTGTGTTACTTGCCATCTAAGGAAGAAAGGTGGCAATATTACACAGCAAGTTTCAG 1575
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1272 TAAAGAGCGGAATATTATACGTGAGATGTTCCGGTCTGGTCTGCCCCCTCTATCTCT 1331
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1576 CGACTACGCCAAGTACTATGCACTTGTCTGCT----- 1607
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1332 ACACAGCGCTGTAATGATAAGGGCTGAGAGTCTCGGAACAAATTCAGCTTTGGATAA 1391
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1608 -----ACGAAATTAATAATCTCGAGAAACCAAGGAATTTGGAAA 1647
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Oy 1392 AATGCTGCAGAAATGTCAGATGCGCTCCAAAAAATCTGGACTTCAATTTATTTTGAATGAAC 1451
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1648 TGCTTTGAAAAATATCCAGCTGCTTAAAGAGGAAATTAAGAACTTGAAGTAGATGAAT 1707
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1452 AAAATTTTGGTATCAGATGATCTTGCCTCTCATTTTGTATTAATCCCAAGAAATATCTCT 1511
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1708 TACTTTATGGTACCAAGATGATTTCTCTCTCAATTTTGACAGATCAAAAGAAATATCTCT 1767
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1512 ACTATTAGATGCTGATGCA----GGCCCATGTAGTCAAAAAAGCAGACACTGTCTTCAGAC 1567
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1768 GCTAATTCAGTGTATGATGGTGGTCCCTGCTCAGTCAGAGTGTAAAGTCTGTATTTGCTG 1827
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1568 TGAATCTGGGCCACTTACTCTTGCAAGCACAGAAAAATTTATATAGTACTAGCTTTTGAATGCA 1627
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1828 TTAATTCGATATCTTATCTTGCAAGTAAAGGAGGTGCTCAATGCTCTGCTGGTGGATGGTC 1887
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1628 GAGGAAGTGGTTACCAAGGAGATAAGATCATGATGCATCAACAGAGACTGGGAAAT 1687
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1888 GAGGAACAGCTTTCCAAAGGTGACAACTCTCTATGCAAGTATTCGAAAGCTGGGTGTTT 1947
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1688 TTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTCAAAAATGGGATTTTGGGACA 1747
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1948 ATGAAGTTGAAGACCAAGATTACAGCTGTCAAGAAATTTCAAGAAATGGGTTTCATTGATG 2007
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1748 ACAAACGAATTTGCAATTTGGGGCTGGTCAATATGGAGGTGAGTAACTCAATGGTCTCTGG 1807
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 2008 AAAAAAGATAGCCATATGGGCTGGTCTTATGAGGAGATAGTTTTCATCACTGGCCCTTG 2067
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1808 GATCGGAAGTGGGCTGTTCAAGTGTGGAATAGCCGTGGCCCTGTATCCCGGTGGGAGT 1867
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 2068 CATCTGGAACCTGGTCTTTTCAAAATGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAAT 2127
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1668 ACTATGACTCAGTGTACACAGAGCTTACATGGTCTGCCAACTCCAGAGACAACTTG 1927
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 2128 ATTAGGGTCTGCTACACAGAGATTCATGGGTCTCCCAAGAGGATGATATCTTG 2187
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1928 ACCATTACAGAAATTCACAGTCTAGCAGCAGAGCTGAAAAATTTTAAACAAAGTTGAGTACC 1987
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 2188 AGCACTATAAGATTTCACTGTGATGCAAGAGCAGAAATATTTTCAGAAATCTAGACTATC 2247
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1988 TCCTTATTCAGAACAGCAGATGATAAGCTTCACTTTTCAGCAGTCTAGCTCAGATCTCCA 2047
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 2248 TTTCTATCCACGGAACAGCAGATGATAATGTGCACTTTTCAAACTCAGCACAGATTGCTA 2307
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 2048 AAGCCCTGGTCTGATTTGGAGTGGATTTCCAGGCAATGTGTATCTGATGAAGACCAATG 2107
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 2308 AAGCTCTGGTTAATGCAAGTGGATTTCCAGGCAATGTGTACTCTGACCAAGAACCAAG 2367
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 2108 GAATAGCTAGCAGCAGCAGCACCAACATATATATATATATATATATATATATATATATAT 2167
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 2368 GCTTATCCGGCTGTCCACGAAC---CACTTATACACCCACATGACCCACTTCTCTAAGC 2424
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 2168 AATGTTTCTCTTT 2180
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 2425 AGTGTCTCTCTTT 2437
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 12

```
US-09-389-681-428
; Sequence 428, Application US/09389681A
; Patent No. 6518237
; GENERAL INFORMATION:
; APPLICANT: Yucui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121-470C3
; CURRENT APPLICATION NUMBER: US/09/389,681A
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: Fast-Seq for Windows Version 3.0
```

```
; SEQ ID NO 428
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-389-681-428

Query Match      12.1%; Score 264.8; DB 3; Length 535;
Best Local Similarity 68.6%; Pred. No. 2.3e-70;
Matches 365; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 1466 AGATGATCTGGCTCCTCAATTTGATAAATCCAGAAATATCCTCTACTATTAGATGTGT 1525
DB 4 AGATGATCTTCTCCTCAATTTGACAGATCAAGAAAGTATCCCTTGTCTTAATTTCAAGTGT 63

QY 1526 ATGCAGGCCCATGTAGTCAAAAGCAGACACTGTCTTCAGACTGAACCTGGGCCACTTACC 1585
DB 64 ATGCTGGTCCCTGATGAGTGAAGTGTAAAGTCTGTATTTGCTGTAAATGGATATCTTATC 123

QY 1586 TTGCAAGCACAGAAACATTATAGTAGTACGTTTCAAAATGGGATGGTTCGAGGAAACAGCTTTCCAAG 183
DB 124 TTGCAAGTAAAGGAGGATGGTCAATGGTGGTGGTTCGAGGAAACAGCTTTCCAAG 183

QY 1646 GAGATAAGATCATGATGATCAACAGAGACTGGGAAACATTTGAAAGTTGAAGATCAAA 1705
DB 184 GTGCAAAACTCCTCTATGACAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAGA 243

QY 1706 TTGAAGCAGCAGACAAATTTTCAAAATGGGATTTGTGCAACAAACAAATTCGAATTT 1765
DB 244 TTACAGCTGTGAGAAATTTCAAGAAATGGGTTTCATGTGTAAGAAAGATAGCCATAT 303

QY 1766 GGGGCTGGTCTATAGGAGGTACGTAACCTCAATGGTCTGGATCGGGAAGTGGCGTGT 1825
DB 304 GGGGCTGGTCTATGAGAGGATACGTTTCATCACTGGCCCTTGCATCTGGAACTGGTCTTT 363

QY 1826 TCAAGTGTGGAATAGCCCTGGGCTGTATCCCGTGGGAGTACTATGACTCAGTGTACA 1885
DB 364 TCAATGTGTATAGTGGTCTCCAGTCTCCAGCTGGGAATATTTACGGCTCTGTCTACA 423

QY 1886 CAGAACGTTACATGGTCTCCCAACTCCAGAGACAACTTGACCATTCAGAAATTCNA 1945
DB 424 CAGAGAGATTCATGGTCTCCCAACTCCAGAGATGATAATCTTGAGCACTATAAGAAATTCNA 483

QY 1946 CAGTCTATGAGCAGAGCTGAAAATTTTAAACAAGTTGAGTACTCTCTTATCA 1997
DB 484 CTGTGATGCGAAGAGCAGAAATATTTTCAGAAATGTAGACTATCTTCTCATCCA 535

RESULT 13
US-09-620-405B-428
; Sequence 428, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 428
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-620-405B-428

Query Match      12.1%; Score 264.8; DB 3; Length 535;
Best Local Similarity 68.6%; Pred. No. 2.3e-70;
Matches 365; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 1466 AGATGATCTGGCTCCTCAATTTGATAAATCCAGAAATATCCTCTACTATTAGATGTGT 1525
DB 4 AGATGATCTTCTCCTCAATTTGACAGATCAAGAAAGTATCCCTTGTCTTAATTTCAAGTGT 63

QY 1526 ATGCAGGCCCATGTAGTCAAAAGCAGACACTGTCTTCAGACTGAACCTGGGCCACTTACC 1585
DB 64 ATGCTGGTCCCTGATGAGTGAAGTGTAAAGTCTGTATTTGCTGTAAATGGATATCTTATC 123

QY 1586 TTGCAAGCACAGAAACATTATAGTAGTACGTTTCAAAATGGGATGGTTCGAGGAAACAGCTTTCCAAG 183
DB 124 TTGCAAGTAAAGGAGGATGGTCAATGGTGGTGGTTCGAGGAAACAGCTTTCCAAG 183

QY 1646 GAGATAAGATCATGATGATCAACAGAGACTGGGAAACATTTGAAAGTTGAAGATCAAA 1705
DB 184 GTGCAAAACTCCTCTATGACAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAGA 243

QY 1706 TTGAAGCAGCAGACAAATTTTCAAAATGGGATTTGTGCAACAAACAAATTCGAATTT 1765
DB 244 TTACAGCTGTGAGAAATTTCAAGAAATGGGTTTCATGTGTAAGAAAGATAGCCATAT 303

QY 1766 GGGGCTGGTCTATAGGAGGTACGTAACCTCAATGGTCTGGATCGGGAAGTGGCGTGT 1825
DB 304 GGGGCTGGTCTATGAGAGGATACGTTTCATCACTGGCCCTTGCATCTGGAACTGGTCTTT 363

QY 1826 TCAAGTGTGGAATAGCCCTGGGCTGTATCCCGTGGGAGTACTATGACTCAGTGTACA 1885
DB 364 TCAATGTGTATAGTGGTCTCCAGTCTCCAGCTGGGAATATTTACGGCTCTGTCTACA 423

QY 1886 CAGAACGTTACATGGTCTCCCAACTCCAGAGACAACTTGACCATTCAGAAATTCNA 1945
DB 424 CAGAGAGATTCATGGTCTCCCAACTCCAGAGATGATAATCTTGAGCACTATAAGAAATTCNA 483

QY 1946 CAGTCTATGAGCAGAGCTGAAAATTTTAAACAAGTTGAGTACTCTCTTATCA 1997
DB 484 CTGTGATGCGAAGAGCAGAAATATTTTCAGAAATGTAGACTATCTTCTCATCCA 535

RESULT 14
US-09-433-826B-428
; Sequence 428, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 428
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-433-826B-428

Query Match      12.1%; Score 264.8; DB 3; Length 535;
Best Local Similarity 68.6%; Pred. No. 2.3e-70;
Matches 365; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 1466 AGATGATCTGGCTCCTCAATTTGATAAATCCAGAAATATCCTCTACTATTAGATGTGT 1525
DB 4 AGATGATCTTCTCCTCAATTTGACAGATCAAGAAAGTATCCCTTGTCTTAATTTCAAGTGT 63

QY 1526 ATGCAGGCCCATGTAGTCAAAAGCAGACACTGTCTTCAGACTGAACCTGGGCCACTTACC 1585
DB 64 ATGCTGGTCCCTGATGAGTGAAGTGTAAAGTCTGTATTTGCTGTAAATGGATATCTTATC 123
```

```

QY 1586 TTGCAAGCAGCAGAAACATTATAGTAGCTAGCTTTGATGGCAGAGAAAGTGGTACCAG 1645
DB 124 TTGCAAGTAAGGAAGGATGGTCAATGGCTTGGTGGATGGTGCAGGAACAGCTTTCCAG 183
QY 1646 GAGATAAGATCATGATGCAATCAACAGAGACTGGGAACATTTGGAAGTTGAAGATCAA 1705
DB 184 GTGACAAACTCCTCTATGCAAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAGA 243
QY 1706 TTGAAGCAGCCAGACAAATTTTCAAAATGGGATTTGTGCAACAAACGAATTCGAATTT 1765
DB 244 TTACAGCTGTTCAGAAATTCATAGAAATGGGTTTCATTTGATGAAAAAAGATAGCCATAT 303
QY 1766 GGGCTGTGTCATATAGGAGGTACGTAACCTCAATGGTCTCCGGATCGGGAAGTGGCGTGT 1825
DB 304 GGGGCTGGTCTCTATGGAGATAGCTTTTCATCACTGGCCCTTGCATCTGGAAGTGGTCTTT 363
QY 1826 TCAAGTGTGGAATAGCCGTGGGCTGTATCCCGTGGGAGTACTATGACTCAGTGTACA 1885
DB 364 TCAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTTACCGCTCTGTCTACA 423
QY 1886 CAGAACGTTTACATGGGCTCTCCAACTCCAGAAAGCAACCTTGACCATTACAGAAATTCAA 1945
DB 424 CAGAGAGATTTCATGGGCTCTCCAAAGAGATGATATCTTTGAGCACTATAGAAATTCAA 483
QY 1946 CAGTCATGAGCAGAGCTGAAATTTTAAACAAAGTTGAGTACCTCTCTATTCA 1997
DB 484 CTGTGATGGCAAGAGCAGAAATTTTCAGAAATGTAGACTATCTTCTCATCCA 535

RESULT 15
US-09-604-287A-428
; Sequence 428, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 428
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-604-287A-428

```

```

Query Match      12.1%; Score 264.8; DB 3; Length 535;
Best Local Similarity 68.6%; Pred. No. 2.3e-70;
Matches 365; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 1466 AGATGATCTTGCCTCTCTCTATTTGATATAATCCAGAAATATCCCTCTACTATTAGATGTT 1525
DB 4 AGATGATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 63
QY 1526 ATGCAGGCCCATGTAGTCAAAAGCAGACACTGTCTTCAGACTGAACTGGGCCACTTACC 1585
DB 64 ATGGTGTCTCTGAGTCAGAGTGTAGTCTGTATTTGCTGTTAATTTGGATATCTTATC 123
QY 1586 TTGCAAGCAGCAAAACATTATAGTAGCTTTTGTATGGCAGAGGAGTGGTTTACCAG 1645
DB 124 TTGCAAGTAAGGAAGGATGGTCAATTTGCTTGGTGGATGGTGCAGGAACAGCTTTTCCAG 183
QY 1646 GAGATAAGATCATGATGCAATCAACAGAGACTGGGAACATTTGGAAGTTGAAGATCAA 1705
DB 184 GTGACAAACTCCTCTATGCAAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAGA 243

```

Search completed: February 17, 2006, 07:48:59
Job time : 406 secs

```

QY 1706 TTGAAGCAGCCAGACAAATTTTCAAAATGGGATTTGTGCAACAAACGAATTCGAATTT 1765
DB 244 TTACAGCTGTTCAGAAATTCATAGAAATGGGTTTCATTTGATGAAAAAAGATAGCCATAT 303
QY 1766 GGGCTGTGTCATATAGGAGGTACGTAACCTCAATGGTCTCCGGATCGGGAAGTGGCGTGT 1825
DB 304 GGGGCTGGTCTCTATGGAGATAGCTTTTCATCACTGGCCCTTGCATCTGGAAGTGGTCTTT 363
QY 1826 TCAAGTGTGGAATAGCCGTGGGCTGTATCCCGTGGGAGTACTATGACTCAGTGTACA 1885
DB 364 TCAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTTACCGCTCTGTCTACA 423
QY 1886 CAGAACGTTTACATGGGCTCTCCAACTCCAGAAAGCAACCTTGACCATTACAGAAATTCAA 1945
DB 424 CAGAGAGATTTCATGGGCTCTCCAAAGAGATGATATCTTTGAGCACTATAGAAATTCAA 483
QY 1946 CAGTCATGAGCAGAGCTGAAATTTTAAACAAAGTTGAGTACCTCTCTATTCA 1997
DB 484 CTGTGATGGCAAGAGCAGAAATTTTCAGAAATGTAGACTATCTTCTCATCCA 535

```

THIS PAGE BLANK (USPTO)

[illegible]

Qy	1441	TTGAATGAAACAAAAATTTTGGTATCAGATGATCTTGGCTCCTCATTTTGTGATAAATCCAAG	1500
Db	1630	TTGAATGAAACAAAAATTTTGGTATCAGATGATCTTGGCTCCTCATTTTGTGATAAATCCAAG	1689
Qy	1501	AAATATCCTCTACTATTAGATGTGTATGCAGGCCCATGTATGATCAAAAAGCAGACACTGTC	1560
Db	1690	AAATATCCTCTACTATTAGATGTGTATGCAGGCCCATGTATGATCAAAAAGCAGACACTGTC	1749
Qy	1561	TTCAGACTGAATCGGGCCACTTACCTTTGCAAGCACAGAAAAACATATTAGTAGCTAGCTTT	1620
Db	1750	TTCAGACTGAATCGGGCCACTTACCTTTGCAAGCACAGAAAAACATATTAGTAGCTAGCTTT	1809
Qy	1621	GATGGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAACAGAGAAGCTG	1680
Db	1810	GATGGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAACAGAGAAGCTG	1869
Qy	1681	GGAACTTTCAAGTTGNAAGATCAAAATGGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT	1740
Db	1870	GGAACTTTCAAGTTGNAAGATCAAAATGGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT	1929
Qy	1741	GTGGACAAACAAACGAATTTGGGCTGGTTCATATCGAGGGTACGTAACCTCAATG	1800
Db	1930	GTGGACAAACAAACGAATTTGGGCTGGTTCATATCGAGGGTACGTAACCTCAATG	1989
Qy	1801	GTCTCTGGGATCGGGAAAGTGGCGTTCATAGTGTGGAAATAGCCGTGGCGCTGTATCCCGG	1860
Db	1990	GTCTCTGGGATCGGGAAAGTGGCGTTCATAGTGTGGAAATAGCCGTGGCGCTGTATCCCGG	2049
Qy	1861	TGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCACTCCAGAGAC	1920
Db	2050	TGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCACTCCAGAGAC	2109
Qy	1921	AACCTTGACCATTTACAGAAATTCACACAGTCATGAGCAGAGCTGAAAAATTTTAAACAAGTT	1980
Db	2110	AACCTTGACCATTTACAGAAATTCACACAGTCATGAGCAGAGCTGAAAAATTTTAAACAAGTT	2169
Qy	1981	GAGTACCTCTCTATTTCATGGAACAGCAGATGATAACGTTCACTTTACGAGTCAAGCTCAG	2040
Db	2170	GAGTACCTCTCTATTTCATGGAACAGCAGATGATAACGTTCACTTTACGAGTCAAGCTCAG	2229
Qy	2041	ATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCTGATGAA	2100
Db	2230	ATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCTGATGAA	2289
Qy	2101	GACCATGGATAGCTAGCAGACAGACACCAACATATATATACCCACATGAGCCACTTC	2160
Db	2290	GACCATGGATAGCTAGCAGACAGACACCAACATATATATACCCACATGAGCCACTTC	2349
Qy	2161	ATAAAACAATGTTTCTCTTTACCT	2184
Db	2350	ATAAAACAATGTTCTCTTTACCT	2373
RESULT 3			
AAL44351			
ID	AAL44351 standard; DNA; 3407 BP.		
XX	AAL44351;		
XX	24-OCT-2002 (first entry)		
XX	Human dipeptidyl peptidase IV (DPP IV) coding sequence.		
KW	Human; angiogenic condition; angiotensin converting enzyme; ACE; gene;		
KW	vasopressinase inhibitor; dipeptidyl peptidase IV; aminopeptidase P; db;		
KW	DPP IV; aminopeptidase P; APP; hypertension; diabetes; cardiac disease;		
XX	renal disease; enzyme.		
OS	Homo sapiens.		
XX	Key		
FH	Location/Qualifiers		
FT	76..2376		
CDS			

Db 490 ||||| AAAAGGCGAGCTGATTTACAGAGAGAGAGATTCACAAACACACACAGTGGGTCCACATGGTCA 549
Qy 361 CCAGTGGGTCAATAATTCGCATATGTTTGGAAACATGACATTTATGTTTAAATTAAGAACCA 420
Db 550 CCAGTGGGTCAATAATTCGCATATGTTTGGAAACATGACATTTATGTTTAAATTAAGAACCA 609
Qy 421 AATTATCAAGTTACAGAAATCACATGGACGGGAAAGAGATATAATATATAATTAAGAAATA 480
Db 610 AATTATCAAGTTACAGAAATCACATGGACGGGAAAGAGATATAATATATAATTAAGAAATA 669
Qy 481 ACTGACTGGGTTATGAAGAGAGAGTCTTCACTGCTACTCTGCTCTGCTGCTGCTGCTGCTGCT 540
Db 670 ACTGACTGGGTTATGAAGAGAGAGTCTTCACTGCTACTCTGCTCTGCTGCTGCTGCTGCTGCT 729
Qy 541 AACGGCAGCTTTTATAGCATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTTATTAATAC 600
Db 730 AACGGCAGCTTTTATAGCATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTTATTAATAC 789
Qy 601 TCCTTCTACTCTGATGAGTCACTGACGATCCCAAGACTGTACGGGTTCCATATCCAAAG 660
Db 790 TCCTTCTACTCTGATGAGTCACTGACGATCCCAAGACTGTACGGGTTCCATATCCAAAG 849
Qy 661 GCAGAGCTGTAATCCAACTGTAAGTTCCTTGTGTAATACAGACTCTCTCAGCTCA 720
Db 850 GCAGAGCTGTAATCCAACTGTAAGTTCCTTGTGTAATACAGACTCTCTCAGCTCA 909
Qy 721 GTACCAATGCACTTCATACAAATCACTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 780
Db 910 GTACCAATGCACTTCATACAAATCACTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 969
Qy 781 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGCTCAGGAGG 840
Db 970 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGCTCAGGAGG 1029
Qy 841 ATTCAAGACTATTCGCTCATGATATTTGTGACTATGATGATCCAGTGGAGATGGAC 900
Db 1030 ATTCAAGACTATTCGCTCATGATATTTGTGACTATGATGATCCAGTGGAGATGGAC 1089
Qy 901 TGCTTAGTGACGCGCAACACATTTGAATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 1090 TGCTTAGTGACGCGCAACACATTTGAATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1149
Qy 961 CCTTCAGAACCTCAATTTTACCTTTGATGTAATAGCTTCTACAAAGATCATCAGCAATGAA 1020
Db 1150 CCTTCAGAACCTCAATTTTACCTTTGATGTAATAGCTTCTACAAAGATCATCAGCAATGAA 1209
Qy 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAGAGCTGACACATTTATTACA 1080
Db 1210 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAGAGCTGACACATTTATTACA 1269
Qy 1081 AAAGGCACCTGGGAGTCACTCGGATAGAGCTCTAACCAAGTATTTATCTATCTACATTT 1140
Db 1270 AAAGGCACCTGGGAGTCACTCGGATAGAGCTCTAACCAAGTATTTATCTATCTACATTT 1329
Qy 1141 AGTAATGAATATAAGGAATCCAGGAGAGAGGAATCTTTATAAAATCCAACTTTATTGAC 1200
Db 1330 AGTAATGAATATAAGGAATCCAGGAGAGAGGAATCTTTATAAAATCCAACTTTATTGAC 1389
Qy 1201 TATACAAAGTGAATGCTCCTCAGTTGTGAGCTGGAATCCGAAAGGTGTCAGTACTATTCT 1260
Db 1390 TATACAAAGTGAATGCTCCTCAGTTGTGAGCTGGAATCCGAAAGGTGTCAGTACTATTCT 1449
Qy 1261 GTGTCATTCAGTAAGGAGGCAAGTATTTATCAGCTGAGATGTTCCGGTCCCTGGTCTGCCC 1320
Db 1450 GTGTCATTCAGTAAGGAGGCAAGTATTTATCAGCTGAGATGTTCCGGTCCCTGGTCTGCCC 1509
Qy 1321 CTCTATCTCTACACAGCAGCGTCAATGATAAAGGCTGAGAGTCCCTGGAGACAAATTC 1380
Db 1510 CTCTATCTCTACACAGCAGCGTCAATGATAAAGGCTGAGAGTCCCTGGAGACAAATTC 1569
Qy 1381 GCTTTGGATAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAATCGGACTTCATTATT 1440

Db 1570 GCTTTGGATAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAATCGGACTTCATTATT 1629
Qy 1441 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGGTATAAATCCCAAG 1500
Db 1630 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGGTATAAATCCCAAG 1689
Qy 1501 AAATATCCTCTCTATTTAGATGTTATGACAGCCCATGTTAGTCAAAAAAGCAGACACTGTC 1560
Db 1690 AAATATCCTCTCTATTTAGATGTTATGACAGCCCATGTTAGTCAAAAAAGCAGACACTGTC 1749
Qy 1561 TTGACACTGAAGTGGGCCACTTACTTTCAGAGCAGAGAAACATTTATAGTAGTACTGCTTT 1620
Db 1750 TTGACACTGAAGTGGGCCACTTACTTTCAGAGCAGAGAAACATTTATAGTAGTACTGCTTT 1809
Qy 1621 GATGGCAGAGAGAGTGGTTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGAGACTG 1680
Db 1810 GATGGCAGAGAGAGTGGTTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGAGACTG 1869
Qy 1681 GGAAATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAAATGGGATTT 1740
Db 1870 GGAAATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAAATGGGATTT 1929
Qy 1741 GTGACACAAACAAACAAATTTGCAATTTGGGCTGGTCAATATGAGGGTACGTAACCTCAATG 1800
Db 1930 GTGACACAAACAAACAAATTTGCAATTTGGGCTGGTCAATATGAGGGTACGTAACCTCAATG 1989
Qy 1801 GTCTCTGGATCGGGAAGTGGGCTGTTCAAGTGTGGAAATAGCGCTGGCGCTGTATCCCGG 1860
Db 1990 GTCTCTGGATCGGGAAGTGGGCTGTTCAAGTGTGGAAATAGCGCTGGCGCTGTATCCCGG 2049
Qy 1861 TGGGAGTACTATGACTCAGTGTACACAGAACTTACATGCGTCTCCCAACTCCAGAGAC 1920
Db 2050 TGGGAGTACTATGACTCAGTGTACACAGAACTTACATGCGTCTCCCAACTCCAGAGAC 2109
Qy 1921 AACCTTGACCTTTACAGAAATTTCAACAGTCTATGAGCAGAGCTGAAATTTTAAACAAAGTT 1980
Db 2110 AACCTTGACCTTTACAGAAATTTCAACAGTCTATGAGCAGAGCTGAAATTTTAAACAAAGTT 2169
Qy 1981 GAGTACCTCTCTATTTATGAAACAGCAGATGATAACGTTTCACTTTTACAGAGTCACTCAG 2040
Db 2170 GAGTACCTCTCTATTTATGAAACAGCAGATGATAACGTTTCACTTTTACAGAGTCACTCAG 2229
Qy 2041 ATCTCCAAAGCCTTGGTCTGATGTTGGAGTGGATTTCCAGGCAATGTGTGTATATCTGATGAA 2100
Db 2230 ATCTCCAAAGCCTTGGTCTGATGTTGGAGTGGATTTCCAGGCAATGTGTGTATATCTGATGAA 2289
Qy 2101 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160
Db 2290 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2349
Qy 2161 ATAAAAACAATGTTTCTCTTTACCT 2184
Db 2350 ATAAAAACAATGTTTCTCTTTACCT 2373

RESULT 5
ADD14640
ID ADD14640 standard; cdna; 3407 BP.
XX
AC ADD14640;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human src biomarker polymucleotide SEQ ID NO:34.
XX
KW predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytosolic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
OS Homo sapiens.
XX
PN W02003062395-A2.
XX

PD 31-JUL-2003.
 XX 17-JAN-2003; 2003WO-US001981.
 XX 18-JAN-2002; 2002US-0350061P.
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA
 PI Huang P, Fairchild CR, Lee FY, Shaw P;
 XX WPI; 2003-636735/60.
 DR P-PSDB; ADD14045.
 DR
 XX New polynucleotides and polypeptides for predicting the activity of
 PT compounds that interact with protein tyrosine kinases and/or protein
 PT tyrosine kinase pathways.
 XX
 PS Claim 2; SEQ ID NO 34; 139pp; English.
 XX
 CC The present invention describes a predictor set comprising a plurality of
 CC polynucleotides or polypeptides whose expression pattern is predictive of
 CC the response of cells to treatment with a compound that modulates protein
 CC tyrosine kinase activity or members of the protein tyrosine kinase
 CC pathway. Also described: (1) predicting whether a compound is capable of
 CC modulating the activity of cells, comprising obtaining a sample of cells,
 CC determining whether the cells express a plurality of markers, and
 CC correlating the expression of the markers to the compound's ability to
 CC modulate the activity of the cells; (2) a plurality of cell lines for
 CC identifying polynucleotides and polypeptides whose expression levels
 CC correlate with compound sensitivity or resistance of cells associated
 CC with a disease state; and (3) identifying polynucleotides and
 CC polypeptides that predict compound sensitivity or resistance of cells
 CC associated with a disease state, comprising subjecting the plurality of
 CC cell lines to one or more compounds, analysing the expression pattern of
 CC a microarray of polynucleotides or polypeptides, and selecting
 CC polynucleotides or polypeptides that predict the sensitivity or
 CC resistance of cells associated with a disease state by using the
 CC expression pattern of the microarray. The polynucleotides and
 CC polypeptides have cytoskeletal activities, and can be used in gene therapy.
 CC The polynucleotides and polypeptides are useful in predicting the
 CC activity of compounds that interact with protein tyrosine kinases and/or
 CC protein tyrosine kinase pathways. These may be used in determining drug
 CC sensitivity in patients to allow the development of individualized
 CC genetic profiles which aid in treating diseases and disorders (e.g.
 CC cancer) based on patient response at a molecular level. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 SQ Sequence 3407 BP; 1077 A; 678 C; 704 G; 948 T; 0 U; 0 Other;
 Query Match 100.0%; Score 2184; DB 10; Length 3407;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGTCGCAAACTTACACTCTAAGTATCTTAAATAAATCTATAGACTGAAGTTATAC 60
 DB 190 AGTCGCAAACTTACACTCTAAGTATCTTAAATAAATCTATAGACTGAAGTTATAC 249
 QY 61 TCCTTAAGTGGATTTGAGATCATGAATATCTCTACAAACAAAGAAATATCTTGGA 120
 DB 250 TCCTTAAGTGGATTTGAGATCATGAATATCTCTACAAACAAAGAAATATCTTGGA 309
 QY 121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGATTT 180
 DB 310 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGATTT 369
 QY 181 GGACATTTCTATCAATGATTTTCAATATCTCTGATGGCAGTTTATTTCTTTAGAATAC 240
 DB 370 GGACATTTCTATCAATGATTTTCAATATCTCTGATGGCAGTTTATTTCTTTAGATAC 429
 QY 241 AACTACGTTGAAGCAATGGAGGCAATCTCTACAGCTTCTATGACATTTATGATTTAAAT 300
 DB 430 AACTACGTTGAAGCAATGGAGGCAATCTCTACAGCTTCTATGACATTTATGATTTAAAT 489

QY 301 AAAAGGAGCTGATTTACAGAAAGAGAGGATTTCCAAACAAACACACAGTGGGTCA 360
 DB 490 AAAAGGAGCTGATTTACAGAAAGAGAGGATTTCCAAACAAACACACAGTGGGTCA 549
 QY 361 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACATGACATTTATGTTTAAATTTGA 420
 DB 550 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACATGACATTTATGTTTAAATTTGA 609
 QY 421 AATTTACCAAGTTACAGAAATCACATGGACGGGGAAGAGATATATATATATATGGAATA 480
 DB 610 AATTTACCAAGTTACAGAAATCACATGGACGGGGAAGAGATATATATATATGGAATA 669
 QY 481 ACTGACTGGGTTTATGAAGAGAGAGTCTTCACTGCTCTCTGCTCTGCTGGTCTTCCA 540
 DB 670 ACTGACTGGGTTTATGAAGAGAGAGTCTTCACTGCTCTCTGCTCTGCTGGTCTTCCA 729
 QY 541 AACGCGACTTTTTTATGACATATGCCCAATTTAAACGACACAGAAAGTCCCACTTTATG 600
 DB 730 AACGCGACTTTTTTATGACATATGCCCAATTTAAACGACACAGAAAGTCCCACTTTATG 789
 QY 601 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 560
 DB 790 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 849
 QY 661 GCAGGAGCTGTGAATCCCAACTGTAAAGTCTTTTGTGTAATACAGAGCTCTCTCAGCTCA 720
 DB 850 GCAGGAGCTGTGAATCCCAACTGTAAAGTCTTTTGTGTAATACAGAGCTCTCTCAGCTCA 909
 QY 721 GTCCCAATGCAACTTCCATACAAATCACTGCTCTCTGCTCTCTATGTTAGAGGATCAC 780
 DB 910 GTCCCAATGCAACTTCCATACAAATCACTGCTCTCTGCTCTCTATGTTAGAGGATCAC 969
 QY 781 TACTTGTGTGATGACATGAGGCAACACAGAAAGAAATTTCTTTCAGTGGTCTCAGGAG 840
 DB 970 TACTTGTGTGATGACATGAGGCAACACAGAAAGAAATTTCTTTCAGTGGTCTCAGGAG 1029
 QY 841 ATTCAAGACTATTCGGTCACTGATGATTTGTGATGATGATGATGATGATGATGATGATG 900
 DB 1030 ATTCAAGACTATTCGGTCACTGATGATTTGTGATGATGATGATGATGATGATGATGATG 1089
 QY 901 TGCTTAGTGGCAACGCAACACATTTGAAATGAGTACTACTGCTGGTGGTGGAAAGATTTAG 960
 DB 1090 TGCTTAGTGGCAACGCAACACATTTGAAATGAGTACTACTGCTGGTGGTGGAAAGATTTAG 1149
 QY 961 CTTTCAGAACCTCATTTTACCTTTCAGTGTGATGATGATGATGATGATGATGATGATGATG 1020
 DB 1150 CTTTCAGAACCTCATTTTACCTTTCAGTGTGATGATGATGATGATGATGATGATGATG 1209
 QY 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAGACTGCACATTTATTACA 1080
 DB 1210 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAGACTGCACATTTATTACA 1269
 QY 1081 AAAGCACCCTGGGAAGTCACTGGGATAGAAGCTCTTAAACAGTGTATCTATATCTATCT 1140
 DB 1270 AAAGCACCCTGGGAAGTCACTGGGATAGAAGCTCTTAAACAGTGTATCTATATCTATCT 1329
 QY 1141 AGTAATGATATATAAGGAATGCCAGGAGAGGAATCTTTATAAATCCAACTTTATGAC 1200
 DB 1330 AGTAATGATATATAAGGAATGCCAGGAGAGGAATCTTTATAAATCCAACTTTATGAC 1389
 QY 1201 TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGGTTGTCAGTACTATTCT 1260
 DB 1390 TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGGTTGTCAGTACTATTCT 1449
 QY 1261 GTGTCATTTCACTAAAGGCGGAAGTATTTATCAGCTGAGATGTTCCGGTCTCGTCCGCC 1320
 DB 1450 GTGTCATTTCACTAAAGGCGGAAGTATTTATCAGCTGAGATGTTCCGGTCTCGTCCGCC 1509
 QY 1321 CTCTATCTCTACACAGCAGCGTGAATGATATAAGGCTGAGAGTCTCTGGAACACATTTCA 1380
 DB 1510 CTCTATCTCTACACAGCAGCGTGAATGATATAAGGCTGAGAGTCTCTGGAACACATTTCA 1569
 QY 1381 GCTTTGGGATAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAACTGGACTTCTATTAT 1440

Qy	121	TTCAATGCTGAATATGGAAACAGCTCAGTTTTTCTTGGAGAACAGTACATTTGATGAGTTT	180
Db	310	TTCAATGCTGAATATGGAAACAGCTCAGTTTTTCTTGGAGAACAGTACATTTGATGAGTTT	369
Qy	181	GGACATTCCTATCAATGATTTAATCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC	240
Db	370	GGACATTCCTATCAATGATTTAATCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC	429
Qy	241	AATACGTGAAGCAATGAGGCATTTCTTACACAGCTTCATATGACATTTATGATTTAAAT	300
Db	430	AATACGTGAAGCAATGAGGCATTTCTTACACAGCTTCATATGACATTTATGATTTAAAT	489
Qy	301	AAAAGGCAGCTGATTTACAGAAAGAGAGGATTCNAAACAACACACAGTGGGTACATGTGCA	360
Db	490	AAAAGGCAGCTGATTTACAGAAAGAGAGGATTCNAAACAACACACAGTGGGTACATGTGCA	549
Qy	361	CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAATTTGAACCA	420
Db	550	CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAATTTGAACCA	609
Qy	421	AATTTACCAAGCTACAGAAATCACATGGACGGGGAAGAAAGATATAAATATATAATGGAATA	480
Db	610	AATTTACCAAGCTACAGAAATCACATGGACGGGGAAGAAAGATATAAATATGGAATA	669
Qy	481	ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGCCCTACTCTGCTCTGTGGTGGTCTCCA	540
Db	670	ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGCCCTACTCTGCTCTGTGGTGGTCTCCA	729
Qy	541	AACGGCATTCTTTTAGCATATGCCCAATTTAAACCAACACAGAAGTCCCACTTATTTGAATAC	600
Db	730	AACGGCATTCTTTTAGCATATGCCCAATTTAAACCAACACAGAAGTCCCACTTATTTGAATAC	789
Qy	601	TCCTTTCTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTGACGGGTTCATATCCAAAG	660
Db	790	TCCTTTCTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTGACGGGTTCATATCCAAAG	849
Qy	661	GCAGGAGCTGGAATCCAACTGTGTAAGTTCCTTTGTTGTAATAACAGACTCTCTCAGCTCA	720
Db	850	GCAGGAGCTGGAATCCAACTGTGTAAGTTCCTTTGTTGTAATAACAGACTCTCTCAGCTCA	909
Qy	721	GTCAACCAATGCAACTTCCTCATACAATCACTGCTCTCTCTATGTTGATAGGGGATCAC	780
Db	910	GTCAACCAATGCAACTTCCTCATACAATCACTGCTCTCTCTCTATGTTGATAGGGGATCAC	969
Qy	781	TACTTGTGTGATGTGACATGGGCAACAACAAGAAAGAAATTTCTTTGCAGTGGCTCAGGAGG	840
Db	970	TACTTGTGTGATGTGACATGGGCAACAACAAGAAAGAAATTTCTTTGCAGTGGCTCAGGAGG	1029
Qy	841	ATTGAGAACTATTGGTCAATGGATATTTGTGACTATGATGATGATGATGATGATGATGATGAT	900
Db	1030	ATTGAGAACTATTGGTCAATGGATATTTGTGACTATGATGATGATGATGATGATGATGATGAT	1089
Qy	901	TGCTTTAGTGGCACGGCAACAATTTGAAATGAGTACTACTGGCTGGGTGGAAAGATTTAGG	960
Db	1090	TGCTTTAGTGGCACGGCAACAATTTGAAATGAGTACTACTGGCTGGGTGGAAAGATTTAGG	1149
Qy	961	CCTTTCAAGAACTCATTTTACCCCTTGATGGTAAATAGCTTCTCAAGATCATCAGCAATGAA	1020
Db	1150	CCTTTCAAGAACTCATTTTACCCCTTGATGGTAAATAGCTTCTCAAGATCATCAGCAATGAA	1209
Qy	1021	GAAGGTTACAGACACATTTGCTATTTCCNATAGATATAAAGACTGCAATTTATTACA	1080
Db	1210	GAAGGTTACAGACACATTTGCTATTTCCNATAGATATAAAGACTGCAATTTATTACA	1269
Qy	1081	AAAGGCACCTGGGAAGTCATCGGGATAGAAGCTCTAAACAGTGATTTATCTATCTACATTT	1140
Db	1270	AAAGGCACCTGGGAAGTCATCGGGATAGAAGCTCTAAACAGTGATTTATCTATCTACATTT	1329
Qy	1141	AGTAATGNAATAAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATTGAC	1200
Db	1330	AGTAATGNAATAAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATTGAC	1389

RESULT 7
ADO19399
ID ADO
XX

AC ADO19399;
 DT 12-AUG-2004 (first entry)
 DE Human PRO polynucleotide #165.
 XX
 KW Human; PRO; gene; 98; immune related disorder;
 KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
 KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
 KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
 KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
 KW renal disease; demyelinating disease; central nervous system;
 KW peripheral nervous system; demyelinating polyneuropathy;
 KW Guillain-Barre syndrome;
 KW chronic inflammatory demyelinating polyneuropathy.
 XX
 OS Homo sapiens.
 XX
 PN WO2004043361-A2.
 XX
 PD 27-MAY-2004.
 XX
 PF 06-NOV-2003; 2003WO-US035268.
 XX
 PR 08-NOV-2002; 2002US-0425235P.
 XX
 XX (GETH) GENENTECH INC.
 PA
 XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
 PI Wood WI, Wu TD;
 FI
 XX
 DR WPI; 2004-420067/39.
 DR P-PSDB; ADO19400.
 XX
 PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
 PT treating an immune related disorder such as systemic lupus erythematosus,
 PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
 PT spondyloarthritis.
 XX
 PS Claim 1; SEQ ID NO 342; 1731pp; English.
 XX
 CC The invention relates to human PRO polypeptides and the polynucleotides
 CC encoding them. The polypeptides and polynucleotides are useful for
 CC treating and diagnosing immune related disorders in mammals. The immune
 CC related disorders include systemic lupus erythematosus, rheumatoid
 CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
 CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
 CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
 CC mellitus, immune-mediated renal disease, demyelinating diseases of the
 CC central or peripheral nervous system, demyelinating polyneuropathy,
 CC Guillain-Barre syndrome and chronic inflammatory demyelinating
 CC polyneuropathy. This sequence represents a human PRO polynucleotide of
 CC the invention.
 XX
 SQ Sequence 3407 BP; 1077 A; 678 C; 704 G; 948 T; 0 U; 0 Other;
 Query Match 100.0%; Score 2184; DB 12; Length 3407;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATACTTATAGACTGAAGTTATAC 60
 Db 190 AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATACTTATAGACTGAAGTTATAC 249
 Qy 61 TCCTTAGATGGATTTCAGATCATGAATATCTTACAAACAGAAAAATAATATCTTGGTA 120
 Db 250 TCCTTAGATGGATTTCAGATCATGAATATCTTACAAACAGAAAAATAATATCTTGGTA 309
 Qy 121 TTCATCTCTGAATATGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
 Db 310 TTCATCTCTGAATATGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 369
 Qy 181 GGACATTTCTATCAATGATTAATCAATATCTCTGATGGCAGTTTATTTCTTCTTGAATAC 240

Db 370 GGACATTTCTATCAATGATTAATCAATATCTCTGATGGCAGTTTATTTCTTGAATAC 429
 Qy 241 AACTACGTGAGCAATGGAGCATTTCTTACAGCTTCATATGATGATTAATTAAT 300
 Db 430 AACTACGTGAGCAATGGAGCATTTCTTACAGCTTCATATGATGATTAATTAAT 489
 Qy 301 AAAAGGCACTGATTTACAGAGAGAGGATTTCCAAACACACACAGTGGGTCACTGGTCA 360
 Db 490 AAAAGGCACTGATTTACAGAGAGAGGATTTCCAAACACACACAGTGGGTCACTGGTCA 549
 Qy 361 CCAGTGGGTCAATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATTGA 420
 Db 550 CCAGTGGGTCAATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATTGA 609
 Qy 421 AATTTACCAAGTTACAGATTCACATGACGCGGGAAGAGATATAATATATATGGAATA 480
 Db 610 AATTTACCAAGTTACAGATTCACATGACGCGGGAAGAGATATAATATATGGAATA 669
 Qy 481 ACTGACTGGGTTTATGAAGAGAGGATTTCTTCACTGCTACTCTGCTGCTGCTGCTGCTCA 540
 Db 670 ACTGACTGGGTTTATGAAGAGAGGATTTCTTCACTGCTACTCTGCTGCTGCTGCTGCTCA 729
 Qy 541 AACGGCACTTTTATGATATGCCCCAATTTAAACACACAGAAAGTCCCACTTATGGAATAC 600
 Db 730 AACGGCACTTTTATGATATGCCCCAATTTAAACACACAGAAAGTCCCACTTATGGAATAC 789
 Qy 601 TCCTTCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 Db 790 TCCTTCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
 Qy 661 GCAGGAGCTGTAATCCAACTGTAAGTCTTTCTTGTGTAAATACAGACTCTCTCAGCTCA 720
 Db 850 GCAGGAGCTGTAATCCAACTGTAAGTCTTTCTTGTGTAAATACAGACTCTCTCAGCTCA 909
 Qy 721 GTCAACCAATGCAACTTCCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
 Db 910 GTCAACCAATGCAACTTCCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 969
 Qy 781 TACTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 Db 970 TACTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
 Qy 841 ATTCAAGAACTATTTCGGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 Db 1030 ATTCAAGAACTATTTCGGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
 Qy 901 TGCTTACTGTCACGGCAACACATTTGAATGATGATGATGATGATGATGATGATGATGATGAT 960
 Db 1090 TGCTTACTGTCACGGCAACACATTTGAATGATGATGATGATGATGATGATGATGATGATGAT 1149
 Qy 961 CCTTCAGAACCTCATTTTACCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 Db 1150 CCTTCAGAACCTCATTTTACCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1209
 Qy 1021 GAAGTTTACAGACACATTTGCTATTTCCAAATAGATAAAAAGACTGCACATTTATTTACA 1080
 Db 1210 GAAGTTTACAGACACATTTGCTATTTCCAAATAGATAAAAAGACTGCACATTTATTTACA 1269
 Qy 1081 AAAGGCACTTGGGAAGTCATCGGATAGAAGCTCTTAAACAGTGAATTTATCTATCTACATTT 1140
 Db 1270 AAAGGCACTTGGGAAGTCATCGGATAGAAGCTCTTAAACAGTGAATTTATCTATCTACATTT 1329
 Qy 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAGGATCTTTTATATAATCCCACTTATTTGAC 1200
 Db 1330 AGTAATGAATATAAAGGAATGCCAGGAGGAGGATCTTTTATATAATCCCACTTATTTGAC 1389
 Qy 1201 TATACAAAAGTGACATGCTCAGTTGTGAGCTGAATCCGAAAAGGTGTGAGTACTATTTCT 1260
 Db 1390 TATACAAAAGTGACATGCTCAGTTGTGAGCTGAATCCGAAAAGGTGTGAGTACTATTTCT 1449
 Qy 1261 GTGTCTTCACTTAAAGAGGCAAGTATTTATCAGCTGAGATGTTTCCGGTCTCTGGTCTGCC 1320

Db 1450 GTGTCATTAGTAAAGAGCGAAGTATTATACAGTGTGAGATGTTCCGGTCTCGTCTGCC 1509
Qy 1321 CTCTATATCTTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGACAATTCA 1380
Db 1510 CTCTATATCTTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGACAATTCA 1569
Qy 1381 GCTTTGGGATAAATGCTGCAGAAATGTCAGAGTGCCCTCCAAAAAATCGACATTCATTAT 1440
Db 1570 GCTTTGGGATAAATGCTGCAGAAATGTCAGAGTGCCCTCCAAAAAATCGACATTCATTAT 1629
Qy 1441 TTGAATGAACAAAAATTTTGGTATCAGATGATCTTTGCCCTCTCATTTTGTATAAATCCAAG 1500
Db 1630 TTGAATGAACAAAAATTTTGGTATCAGATGATCTTTGCCCTCTCATTTTGTATAAATCCAAG 1689
Qy 1501 AAATATCTCTACTATTAGATGTGTATGACAGGCCCATGTAGTCAAAAAGCAGACATGTC 1560
Db 1690 AAATATCTCTACTATTAGATGTGTATGACAGGCCCATGTAGTCAAAAAGCAGACATGTC 1749
Qy 1561 TTCAGACTGAACCTGGGCCACTTTACCTTGAAGCAGACAGAAAAATTTATAGTAGTACGTTT 1620
Db 1750 TTCAGACTGAACCTGGGCCACTTTACCTTGAAGCAGACAGAAAAATTTATAGTAGTACGTTT 1809
Qy 1621 GATGCGAGAGGAAGTGGTTTACCAAGGAGATAGATCATGCAATCAACAGAAAGACTG 1680
Db 1810 GATGCGAGAGGAAGTGGTTTACCAAGGAGATAGATCATGCAATCAACAGAAAGACTG 1869
Qy 1681 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1740
Db 1870 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1929
Qy 1741 GTGGACAACAAAGCAATTTGCAATTTGGGGCTGCTCATATGGAGGGTACGTAACCTCAATG 1800
Db 1930 GTGGACAACAAAGCAATTTGCAATTTGGGGCTGCTCATATGGAGGGTACGTAACCTCAATG 1989
Qy 1801 GTCTCGGATCGGGAAGTGGCTGTTTCAAGTGTGGAATAGCGGTGGCGCTGTATCCCGG 1860
Db 1990 GTCTCGGATCGGGAAGTGGCTGTTTCAAGTGTGGAATAGCGGTGGCGCTGTATCCCGG 2049
Qy 1861 TGGGAGTACTATGACTCAGTGTACAGAAAGTTCATAGTGGTCTCCCAACTCCAGAAGAC 1920
Db 2050 TGGGAGTACTATGACTCAGTGTACAGAAAGTTCATAGTGGTCTCCCAACTCCAGAAGAC 2109
Qy 1921 AACCTTGACCATTTACAGAAATTTCAACAGTCATGACAGAGCTGAAATTTTAAACAAGTT 1980
Db 2110 AACCTTGACCATTTACAGAAATTTCAACAGTCATGACAGAGCTGAAATTTTAAACAAGTT 2169
Qy 1981 GAGTACCTCTTATTCATGGAACAGCAGATGATAACGTTTCACTTTTCAGCAGTCAGTCAG 2040
Db 2170 GAGTACCTCTTATTCATGGAACAGCAGATGATAACGTTTCACTTTTCAGCAGTCAGTCAG 2229
Qy 2041 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATGATGAA 2100
Db 2230 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATGATGAA 2289
Qy 2101 GACCATGGAATAGCTAGCAGCAGCAGCAGACCAACCATATATATACCCACATGAGCCACTTC 2160
Db 2290 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2349
Qy 2161 ATAAACCAATGTTCTCTTTACCT 2184
Db 2350 ATAAACCAATGTTCTCTTTACCT 2373

RESULT 8
ADV25524
ID ADV25524 standard; cDNA; 3407 BP.
XX
AC ADV25524;
XX
XX
DT 24-FEB-2005 (first entry)
XX
DE Human dipeptidyl-peptidase IV coding sequence.
XX

KW Dipeptidyl-peptidase IV; DPP4; cardiovascular disease;
KW dermatological disease; cancer; neoplasm; hematological disease;
KW respiratory disease; gastrointestinal disease; liver disease;
KW metabolic disorder; Cardiovascular-Gen.; Endocrine-Gen.;
KW Antiinflammatory; Gastrointestinal-Gen.; Gynecological; Hepatotropic;
KW Neuroprotective; Cystostatic; Antiparkinsonian; Nootropic; Cardiant;
KW Antiarrhythmic; Antiartherosclerotic; Antianemic; Antidiabetic;
KW Dermatological; Immunosuppressive; Muscular-Gen.; Antirheumatic;
KW Antiarthritic; Antipeoriatic; Antiinfertility; Gene Therapy; gene; ss.
XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 76..2376
FT /*tag= a
FT /product= "Human DPP4"
XX
XX WO2004104216-A2.
XX
XX 02-DEC-2004.
XX
XX 12-MAY-2004; 2004WO-EP005071.
XX
XX 21-MAY-2003; 2003EP-00011481.
XX (FARB) BAYER HEALTHCARE AG.
XX
XX Golz S, Brueggemeier U, Summer H;
PI WPI; 2004-834301/82.
DR P-PSDB; ADV25525.
DR GENBANK; NM_001935.
XX
XX Use of dipeptidylpeptidase IV (DPP4) polypeptides or polynucleotides for
PT screening therapeutic agents or for diagnosing or treating diseases
PT associated with DPP4, e.g. cardiovascular, metabolic, inflammatory, or
PT neurological disorders.
XX
XX Disclosure; SEQ ID NO 1; 128pp; English.
PS
XX
XX The present sequence is the coding sequence for human dipeptidyl-
CC peptidase IV (DPP4). The invention relates to novel disease associations
CC of DPP4 polypeptides and polynucleotides and to novel methods of
CC screening for therapeutic agents for the treatment of cardiovascular
CC disorders, dermatological disorders, cancer, hematological disorders,
CC respiratory diseases, gastrointestinal and liver diseases, urological
CC disorders and metabolic diseases. Pharmaceutical compositions are
CC provided for treatment of these diseases and disorders and comprise a
CC DPP4 polypeptide, a DPP4 polynucleotide, or regulators of DPP4 or
CC modulators of DPP4 activity. The therapeutic agent is preferably a small
CC molecule, an RNA molecule, an antisense oligonucleotide, a polypeptide,
CC an antibody or a ribozyme. The invention also provides methods of
CC diagnosing diseases and disorders associated with DPP4 by measuring the
CC amount of a DPP4 polynucleotide in a sample and comparing it with the
CC amount in a sample from a healthy and/or diseased mammal. The diseases
CC and disorders include Parkinson's disease, dementia, Alzheimer's disease,
CC myocardial infarction, arrhythmias, atherosclerosis, anemia, eosinophilic
CC disorders, leukemia, pancreatitis, Crohn's disease, inflammatory bowel
CC disease, diabetes, Cushing's syndrome, systemic lupus erythematosus,
CC myasthenia gravis, rheumatoid arthritis, psoriasis, scleroderma, or
CC infertility.
XX
XX Sequence 3407 BP; 1077 A; 678 C; 704 G; 948 T; 0 U; 0 Other;

Query Match 100.0%; Score 2184; DB 13; Length 3407;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGTCGCAAAACCTTACACTCTAACTTACTTAAAAAATCTATAGACTGAGTTATAC 60
Db 190 AGTCGCAAAACCTTACACTCTAACTTACTTAAAAAATCTATAGACTGAGTTATAC 249
Qy 61 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAGAAAAATAATATCTTGGTA 120

[illegible]

Db	1330	AGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAAATCCAACCTATTATGGAC	1389
Qy	1201	TATACAAAAGTGACATGCCCTCAGTTGTGAGCTGAATCCGGAAGGTGTGTCAGTACTATTCT	1260
Db	1390	TATACAAAAGTGACATGCCCTCAGTTGTGAGCTGAATCCGGAAGGTGTGTCAGTACTATTCT	1449
Qy	1261	GTGTCATTCAGTAAAGAGGGCAAGTATTATCAGCTGAGATGTTTCCGGTCTCGTGTGCCCC	1320
Db	1450	GTGTCAATTCAGTAAAGAGGGCAAGTATTATCAGCTGAGATGTTTCCGGTCTCGTGTGCCCC	1509
Qy	1321	CTCTATACTCTCACACAGCAGCGTAATGATAAAGGCTGAGAGTCTCTGGAAGACAATTCA	1380
Db	1510	CTCTATACTCTCACACAGCAGCGTAATGATAAAGGCTGAGAGTCTCTGGAAGACAATTCA	1569
Qy	1381	GCTTTGGATAAAATGCTGCGAAGATGTCACAGATGCCCTCCAAAAACCTGACATTCATTAT	1440
Db	1570	GCTTTGGATAAAATGCTGCGAAGATGTCACAGATGCCCTCCAAAAACCTGACATTCATTAT	1629
Qy	1441	TTGGAATCAAAACAAAATTTTGGTATCAGATGATCTTGCCCTCCTCATTTTGCATAAATCCAAG	1500
Db	1630	TTGGAATCAAAACAAAATTTTGGTATCAGATGATCTTGCCCTCCTCATTTTGCATAAATCCAAG	1689
Qy	1501	AAATATCCTCTACTATTAGATGTGTATGCAAGGCCCATGTAGTCAAAAAGCAGACACTGTC	1560
Db	1690	AAATATCCTCTACTATTAGATGTGTATGCAAGGCCCATGTAGTCAAAAAGCAGACACTGTC	1749
Qy	1561	TTCAGACTGAATCGGGCCACTTACCTTGCACAGACACAGAAAAATATATAGTAGTACGTTTT	1620
Db	1750	TTCAGACTGAATCGGGCCACTTACCTTGCACAGACACAGAAAAATATATAGTAGTACGTTTT	1809
Qy	1621	GATGGCAGAGGAAGTGTGTACCAAGCAGATTAAGATCATGTCATCAACAGAGACACTG	1680
Db	1810	GATGGCAGAGGAAGTGTGTACCAAGCAGATTAAGATCATGTCATCAACAGAGACACTG	1869
Qy	1681	GGAACTTTTCAAGTTGAAGATCAAATTTGAAGCAGCCAGACAATTTTTCAAAAAATGGGATTT	1740
Db	1870	GGAACTTTTCAAGTTGAAGATCAAATTTGAAGCAGCCAGACAATTTTTCAAAAAATGGGATTT	1929
Qy	1741	GTGGACACAAAACGAATTCGAATTTGGGGCTGGTCATATGGAGGGTACGTAAACCTCAATG	1800
Db	1930	GTGGACACAAAACGAATTCGAATTTGGGGCTGGTCATATGGAGGGTACGTAAACCTCAATG	1989
Qy	1801	GTCTGGGATCGGGAAATGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	1860
Db	1990	GTCTGGGATCGGGAAATGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	2049
Qy	1861	TGGGAGTATATAGCTCAGTGTAACAGAAACGTTTACATGGGTCTCCAACTCCAGAAAC	1920
Db	2050	TGGGAGTATATAGCTCAGTGTAACAGAAACGTTTACATGGGTCTCCAACTCCAGAAAC	2109
Qy	1921	AACCTTGACCATTACAGAAATTCCAAAGTCATGAGCAGAGCTGAAAATTTTAAAAACAAGTT	1980
Db	2110	AACCTTGACCATTACAGAAATTCCAAAGTCATGAGCAGAGCTGAAAATTTTAAAAACAAGTT	2169
Qy	1981	GAGTACCTCTTATTATGSAACAGCAGATGATAACGTTTCACTTTCAGCAGTCAGCTCAG	2040
Db	2170	GAGTACCTCTTATTATGSAACAGCAGATGATAACGTTTCACTTTCAGCAGTCAGCTCAG	2229
Qy	2041	ATCTCCAAAGCCCTGGTTCGATGTTTGGAGTGGATTTCCAGGCAATGTGGTATACTGATGAA	2100
Db	2230	ATCTCCAAAGCCCTGGTTCGATGTTTGGAGTGGATTTCCAGGCAATGTGGTATACTGATGAA	2289
Qy	2101	GACCATGGGAATAGCTAGCAGCACAGCACACCAACATATATATACCCACATGAGCCACTTC	2160
Db	2290	GACCATGGGAATAGCTAGCAGCACAGCACACCAACATATATATACCCACATGAGCCACTTC	2349
Qy	2161	ATAAAAAATATGTTTCTCTTTACTT	2184
Db	2350	ATAAAAAATATGTTTCTCTTTACTT	2373

[illegible]

Db	2155	ATCTCCAAAGCCCTGGTGGATGTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA	2214
Qy	2101	GACCATGGAATAGCTAGCAGCACAGCACACCAACATATATATACCCACATGAGCCACTTC	2160
Db	2215	GACCATGGAATAGCTAGCAGCACAGCACACCAACATATATATACCCACATGAGCCACTTC	2274
Qy	2161	ATAAAACAATGTTTCTCTTTACCT	2184
Db	2275	ATAAAACAATGTTTCTCTTTACCT	2298
RESULT 10			
ADJ83980	ADJ83980 standard; DNA; 2301 BP.		
XX	AC	ADJ83980;	
XX	06-MAY-2004	(first entry)	
XX	Human full-length colon dipeptidyl peptidase IV (DPPIV) DNA.		
KW	crystal; proteni co-ordinate data; dipeptidyl peptidase IV; DPPIV;		
KW	immunomodulatory; antidiabetic; antiinflammatory; neuroprotective;		
KW	antithyroid; antineumatic; antiarthritic; anti-HIV; cytosclastic;		
KW	immune response; diabetes; inflammation; multiple sclerosis;		
KW	Grave's disease; chronic rheumatoid arthritis; AIDS; cancer; human;		
KW	colon; ds; gene.		
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	CDS	1..2301	
FT	/*tag= a		
FT	/product= "Human full-length colon dipeptidyl peptidase		
FT	IV (DPPIV) protein"		
XX			
PN	WO2004011640-A1.		
XX			
PD	05-FEB-2004.		
XX			
PF	28-JUL-2003; 2003WO-JP009523.		
XX			
PR	29-JUL-2002; 2002US-0398761P.		
XX			
PA	(TANA) TANABE SEIVAKU CO.		
XX			
PI	Hiramatsu H, Kyono K, Shima H, Sugiyama S;		
XX			
DR	WPI; 2004-156830/15.		
DR	P-PSDB; ADJ83981.		
XX			
PT	New crystal of dipeptidyl peptidase IV capable of analyzing its three-		
PT	dimensional structure, useful for designing, identifying, evaluating or		
PT	searching an effector of the dipeptidyl peptidase IV.		
XX			
PS	Disclosure; SEQ ID NO 1; 332pp; English.		
XX			
CC	The invention relates to a novel crystal of a dipeptidyl peptidase IV		
CC	(DPPIV) which is sufficient to ensure a resolution capable of analysing		
CC	its three-dimensional structure up to the side chain level by X-ray		
CC	crystallographic structural analysis. The crystal of the invention		
CC	demonstrates immunomodulatory, antidiabetic, antiinflammatory,		
CC	neuroprotective, antithyroid, antineumatic, antiarthritic, anti-HIV and		
CC	cytostatic activities and may be useful for providing a three-dimensional		
CC	structural coordinate as the information for designing, identifying,		
CC	evaluating or searching for an effector of the dipeptidyl peptidase IV.		
CC	The effector may be useful as a modulatory agent of immune response and		
CC	as a therapeutic or prophylactic agent for diabetes, inflammation,		
CC	multiple sclerosis, Grave's disease, chronic rheumatoid arthritis, AIDS		
CC	or cancer. The current sequence is that of the human full-length colon		
CC	dipeptidyl peptidase IV (DPPIV) DNA of the invention.		
XX			
SQ	Sequence 2301 BP; 733 A; 453 C; 493 G; 622 T; 0 U; 0 Other;		

Query Match				99.9%; Score 2182.4; DB 12; Length 2301;			
Best Local Similarity				99.9%; Pred. No. 0;			
Matches 2183; Conservative				0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	AGTCGCAAACTTACACTCTA	CTA	CTA	CTA	CTA	CTA
Db	115	AGTCGCAAACTTACACTCTA	CTA	CTA	CTA	CTA	CTA
Qy	61	TCCTTAAGATGGATTT	CAGAT	CAT	GAAT	CTCT	CAAA
Db	175	TCCTTAAGATGGATTT	CAGAT	CAT	GAAT	CTCT	CAAA
Qy	121	TTCAATTCGTGAATAT	GGAA	CAG	CTCAG	TTTTCT	TTGG
Db	235	TTCAATTCGTGAATAT	GGAA	CAG	CTCAG	TTTTCT	TTGG
Qy	181	GGACATTCATCAATGAT	TATTA	CTAT	CTCT	GAT	GGC
Db	295	GGACATTCATCAATGAT	TATTA	CTAT	CTCT	GAT	GGC
Qy	241	AACCTAGCTGAAGCAAT	GGAGG	CAATTC	CTAC	CAG	CTT
Db	355	AACCTAGCTGAAGCAAT	GGAGG	CAATTC	CTAC	CAG	CTT
Qy	301	AAAAGGCAGCTGATTA	CAGAG	AGGATTC	CAAA	CAAC	AC
Db	415	AAAAGGCAGCTGATTA	CAGAG	AGGATTC	CAAA	CAAC	AC
Qy	361	CCAGTGGGTCATAAAT	TGGC	ATAT	TGTT	GGAA	CA
Db	475	CCAGTGGGTCATAAAT	TGGC	ATAT	TGTT	GGAA	CA
Qy	421	AATTTACCAAGTTTAC	AGAT	TCAT	GGC	GGG	AA
Db	535	AATTTACCAAGTTTAC	AGAT	TCAT	GGC	GGG	AA
Qy	481	ACTGACTGGGTTTAT	CAAG	GAAG	GTCT	CAG	TG
Db	595	ACTGACTGGGTTTAT	CAAG	GAAG	GTCT	CAG	TG
Qy	541	AACGGCACTTTTTAG	CATAT	GCCA	ATTTAA	CGA	CA
Db	655	AACGGCACTTTTTAG	CATAT	GCCA	ATTTAA	CGA	CA
Qy	601	TCCTTCTACTGTGAT	GATC	ACTG	CAGT	CAAC	AG
Db	715	TCCTTCTACTGTGAT	GATC	ACTG	CAGT	CAAC	AG
Qy	661	GCAGGAGCTGTGAAT	CCAA	CTGT	AAAG	TTCTTT	GT
Db	775	GCAGGAGCTGTGAAT	CCAA	CTGT	AAAG	TTCTTT	GT
Qy	721	GTCAACCAATGCAAT	TTCC	ATAC	CAAA	TCAT	CT
Db	835	GTCAACCAATGCAAT	TTCC	ATAC	CAAA	TCAT	CT
Qy	781	TACTTGTGTGATGTG	ACAT	GGG	CAAC	AGAA	GA
Db	895	TACTTGTGTGATGTG	ACAT	GGG	CAAC	AGAA	GA
Qy	841	ATTTCAGAACTATT	TCGGT	CAT	GGAT	ATTTG	TG
Db	955	ATTTCAGAACTATT	TCGGT	CAT	GGAT	ATTTG	TG
Qy	901	TGCTTAGTGCGCAG	CGCA	CA	CAAT	TGAA	T
Db	1015	TGCTTAGTGCGCAG	CGCA	CA	CAAT	TGAA	T
Qy	961	CTTTCAGAACTCAT	TTTTA	CCCT	TGAT	GTGA	T
Db	1075	CTTTCAGAACTCAT	TTTTA	CCCT	TGAT	GTGA	T

Db 1375 GTGTCATTCAAGTAAGAGCGGAAGTATTATCAGTCGAGATGTTCCGGTCTCGTCTGCC 1434
Qy 1321 CTCCTATACCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGCAATTC 1380
Db 1435 CTCCTATACCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGCAATTC 1494
Qy 1381 GCTTTGGATAAATGCTGCAGATGTCAGATGCCCTCCAAAACACTGGACTTCATTATT 1440
Db 1495 GCTTTGGATAAATGCTGCAGATGTCAGATGCCCTCCAAAACACTGGACTTCATTATT 1554
Qy 1441 TTGAATGAAACAAAATTTTGGGTATCAGATGATCTTGCCTCTCTCATTTTGTATATAATCC 1500
Db 1555 TTGAATGAAACAAAATTTTGGGTATCAGATGATCTTGCCTCTCTCATTTTGTATATAATCC 1614
Qy 1501 AAATATCCTCTACTATTAGATGTGTATGAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1560
Db 1615 AAATATCCTCTACTATTAGATGTGTATGAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1674
Qy 1561 TTCAGACTGAACCTGGGCCACTTACCTTGCAGCAGACAGAAACATTTATAGTACTAGCTTT 1620
Db 1675 TTCAGACTGAACCTGGGCCACTTACCTTGCAGCAGACAGAAACATTTATAGTACTAGCTTT 1734
Qy 1621 GATGCGAGAGGAGTGGTTTACCAAGGAGATAGATCATGATGCAATCAACAGAGACTG 1680
Db 1735 GATGCGAGAGGAGTGGTTTACCAAGGAGATAGATCATGATGCAATCAACAGAGACTG 1794
Qy 1681 GGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 1740
Db 1795 GGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 1854
Qy 1741 GTGGCAACAAACGAATTCGAATTTGGGGCTGGTCATATGGAGGGTACGTTAACCCTCAATG 1800
Db 1855 GTGGCAACAAACGAATTCGAATTTGGGGCTGGTCATATGGAGGGTACGTTAACCCTCAATG 1914
Qy 1801 GTCCTGGGATCGGGAAGTGGCGTGTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1860
Db 1915 GTCCTGGGATCGGGAAGTGGCGTGTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1974
Qy 1861 TGGGAGTACTATGACTCAGTGTACACAGAACGTTTACATGGGTCTCCCAACTCCAGAAGAC 1920
Db 1975 TGGGAGTACTATGACTCAGTGTACACAGAACGTTTACATGGGTCTCCCAACTCCAGAAGAC 2034
Qy 1921 AACCTTGACCATTTACAGAAATTCACAGTCATGAGCAGAGCTGAAATTTTAAACAAGTT 1980
Db 2035 AACCTTGACCATTTACAGAAATTCACAGTCATGAGCAGAGCTGAAATTTTAAACAAGTT 2094
Qy 1981 GAGTACCTCTTATTTCATGGAACAGCAGATGATTAAGCTTTCAGCAGTCAGCTCAG 2040
Db 2095 GAGTACCTCTTATTTCATGGAACAGCAGATGATTAAGCTTTCAGCAGTCAGCTCAG 2154
Qy 2041 ATCTCCAAAGCCCTGCTCGATGTTGGAGTGGATTTTCCAGGCAATGTTGATATCTGATGAA 2100
Db 2155 ATCTCCAAAGCCCTGCTCGATGTTGGAGTGGATTTTCCAGGCAATGTTGATATCTGATGAA 2214
Qy 2101 GACCATGGAATAGCTAGCAGCAGCAGACACCAACATATATATACCAATGAGCCACTTC 2160
Db 2215 GACCATGGAATAGCTAGCAGCAGCAGACACCAACATATATATACCAATGAGCCACTTC 2274
Qy 2161 ATAAACAATGTTTCTCTTACCT 2184
Db 2275 ATAAACAATGTTTCTCTTACCT 2298

RESULT 12
ID ADO71643
XX ADO71643 standard; DNA; 2461 BP.
AC ADO71643;
XX
DT 26-AUG-2004 (first entry)
XX Nucleotide sequence of a human CD26 protein.
XX

KW CD26: chemotherapeutic; radiotherapeutic; cancer; cell growth;
KW dipeptidyl peptidase IV; DPPIV; topoisomerase II inhibitor;
KW cell cycle arrest; tumour; tumour necrosis; immune response; human; ss;
KW gene.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 26..2326
XX /*tag= a
XX /product= "CD26"
XX WO2004045497-A2.
XX
XX 03-JUN-2004.
XX
XX 15-MAY-2003; 2003WO-US015499.
XX
XX 17-MAY-2002; 2002US-0381606P.
XX (TEXA) UNIV TEXAS SYSTEM.
XX
XX Dang NH, Morimoto C;
XX
XX WPI: 2004-420511/39.
XX P-PSDB; ADO71644.
XX
XX Use of a CD26 composition, and a chemotherapeutic and/or a
XX radiotherapeutic agent for e.g. inhibiting the cell growth, inducing cell
XX cycle arrest, killing a cancer cell, treating cancer, or inducing tumor
XX regression or tumor necrosis.
XX
XX Claim 14; Page 171-175; 182pp; English.
XX
XX The specification describes a CD26 composition which, in conjunction with
XX chemotherapeutic or radiotherapeutic agents, is used for the treatment
XX and prevention of cancers. Expression of CD26 enhances the sensitivity of
XX the cancer cell to the chemotherapeutic or radiotherapeutic agent. CD26
XX is a dipeptidyl peptidase IV (DPPIV). The chemotherapeutic agent is a
XX topoisomerase II inhibitor. The CD26 composition of the invention is a
XX useful for inhibiting the growth of a cell, inducing cell cycle arrest in
XX a cell, killing a cancer cell, potentiating the effect of a
XX chemotherapeutic agent and/or a radiotherapeutic agent on a tumour cell,
XX inducing or enhancing apoptosis of a cancer cell, treating cancer, or
XX inducing tumour regression or tumour necrosis. The CD26 composition is
XX further useful for increasing topoisomerase II expression in a cell, for
XX activating an antigen-presenting cell, or for potentiating immune
XX responses of an animal. The present sequence encodes a CD26 protein, and
XX is used to create vectors which are used to produce compositions of the
XX invention.

XX
SQ Sequence 2461 BP; 805 A; 484 C; 513 G; 659 T; 0 U; 0 Other;
Query Match 99.9%; Score 2182.4; DB 12; Length 2461;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AGTCGCAAAACCTTACACTCTACTGATTACTTTAAAAAATCTTTAGACTGAAGTTATAC 60
Db 140 AGTCGCAAAACCTTACACTCTACTGATTACTTTAAAAAATCTTTAGACTGAAGTTATAC 199
Qy 61 TCCTTAAGATGGATTTTCAGATCATGATATCTCTCAAAACAGAAATATATCTTGGTA 120
Db 200 TCCTTAAGATGGATTTTCAGATCATGATATCTCTCAAAACAGAAATATATCTTGGTA 259
Qy 121 TTCAATGCTGAATATGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTCATGAGTTT 180
Db 260 TTCAATGCTGAATATGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTCATGAGTTT 319
Qy 181 GGACATCTTATCAATGATTATTTCATATCTCTGATGGGCGAGTTTATCTCTTAGAATAC 240
Db 320 GGACATCTTATCAATGATTATTTCATATCTCTGATGGGCGAGTTTATCTCTTAGAATAC 379

QY 241 AACTACGTGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300
Db AACTACGTGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 439
QY 301 AAAAGGCAAGCTGATTTACAGAGAGAGGATTCCTAACACACACACAGTGGGTACATGCTCA 360
Db AAAAGGCAAGCTGATTTACAGAGAGAGGATTCCTAACACACACACAGTGGGTACATGCTCA 499
QY 361 CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATTTGAACCA 420
Db CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATTTGAACCA 559
QY 421 AATTATCAAGTTACAGAAATCACATGGAAGGGAAGAAAGATATAAATATAAATGGAATA 480
Db AATTATCAAGTTACAGAAATCACATGGAAGGGAAGAAAGATATAAATATAAATGGAATA 619
QY 481 ACTGACTGGGTTTATGAGAGAGAGTCTTCAAGTGCCTACTCTGCTCTGTGGTGGTCTCCA 540
Db ACTGACTGGGTTTATGAGAGAGAGTCTTCAAGTGCCTACTCTGCTCTGTGGTGGTCTCCA 679
QY 541 AACGGCACTTTTATAGCATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTTATTTGAATAC 600
Db AACGGCACTTTTATAGCATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTTATTTGAATAC 739
QY 601 TCCTTCTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACCGGTTCCATATCCAAAG 660
Db TCCTTCTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACCGGTTCCATATCCAAAG 799
QY 661 GCAGGAGCTGGAATCCAACTGTAAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCA 720
Db GCAGGAGCTGGAATCCAACTGTAAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCA 859
QY 721 GTCCACCAATGCAATCTTCATACAAATCACTGCTCCTGCTCTATGTTGATAGGGGATCAC 780
Db GTCCACCAATGCAATCTTCATACAAATCACTGCTCCTGCTCTATGTTGATAGGGGATCAC 919
QY 781 TACTTGTGTGATGTGACATGGGCAACAAGAAGAAATTTCTTTGCACTGAGGCTCAGGAGG 840
Db TACTTGTGTGATGTGACATGGGCAACAAGAAGAAATTTCTTTGCACTGAGGCTCAGGAGG 979
QY 841 ATTGAGAACTATTGGTCAATGATATTTGTGACTATGATGATTCAGTGGAGATGGAAAC 900
Db ATTGAGAACTATTGGTCAATGATATTTGTGACTATGATGATTCAGTGGAGATGGAAAC 1039
QY 901 TGCTTAGTGGCAGCGCAACATTTGAATGAGTACTCTGCTGGGTTGGAGATTTAGG 960
Db TGCTTAGTGGCAGCGCAACATTTGAATGAGTACTCTGCTGGGTTGGAGATTTAGG 1099
QY 961 CTTTCAGAACTCAATTTTACCTTTGATGGTAATAGCTTTCTACAAGATCATCAGCAATGAA 1020
Db CTTTCAGAACTCAATTTTACCTTTGATGGTAATAGCTTTCTACAAGATCATCAGCAATGAA 1159
QY 1021 GAAGGTTACAGACATTTGCTATTTTCCAAATAGATAAATAAGAACTGACATTTATTTACA 1080
Db GAAGGTTACAGACATTTGCTATTTTCCAAATAGATAAATAAGAACTGACATTTATTTACA 1219
QY 1081 AAAGSCACTGGGAGTCACTCGGATAGAAGCTCTAACCCAGTGATTTACTACTACATT 1140
Db AAAGSCACTGGGAGTCACTCGGATAGAAGCTCTAACCCAGTGATTTACTACTACATT 1279
QY 1141 AGTAATGAATATAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTTATTGAC 1200
Db AGTAATGAATATAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTTATTGAC 1339
QY 1201 TATACAAAGTGACATGCTCAGTTGTGAGTGAATCCGGAAGAGGTGTCAGTACTATTCT 1260
Db TATACAAAGTGACATGCTCAGTTGTGAGTGAATCCGGAAGAGGTGTCAGTACTATTCT 1399
QY 1261 GTGTCAATTCAGTAAAGGCGGAATTTATCAGCTGAGATGTTCCGGTCTGCTGGTCCGCC 1320
Db GTGTCAATTCAGTAAAGGCGGAATTTATCAGCTGAGATGTTCCGGTCTGCTGGTCCGCC 1459
QY 1321 CTCTATACTCTACACAGCAGCGGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGACAAATTTCA 1380

Db 1460 CTCTATATCTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGACAAATTTCA 1519
QY 1381 GCTTTGGATATAAATGCTTGCAGAAATGTCAGATGCCCTCCAAAAAATCGACTTCATTTATTT 1440
Db 1520 GCTTTGGATATAAATGCTTGCAGAAATGTCAGATGCCCTCCAAAAAATCGACTTCATTTATTT 1579
QY 1441 TTGAATGAAAAAATAAATTTTGGTATCAGATGATCTTGGCTCTCTCAATTTTGTATAAATCCAAAG 1500
Db 1580 TTGAATGAAAAAATAAATTTTGGTATCAGATGATCTTGGCTCTCTCAATTTTGTATAAATCCAAAG 1639
QY 1501 AAATATCTCTACTATTTAGATGTTGATGTCAGGCCCATGTCAGTCAAAAAGACGACACTGTC 1560
Db 1640 AAATATCTCTACTATTTAGATGTTGATGTCAGGCCCATGTCAGTCAAAAAGACGACACTGTC 1699
QY 1561 TTCAAGACTGAACTCGGGCCACTTACCTTTGCAAGCAGACAGAAAAATTTATAGTAGCTAGCTTT 1620
Db 1700 TTCAAGACTGAACTCGGGCCACTTACCTTTGCAAGCAGACAGAAAAATTTATAGTAGCTAGCTTT 1759
QY 1621 GATGCGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGACTG 1680
Db 1760 GATGCGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGACTG 1819
QY 1681 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1740
Db 1820 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1879
QY 1741 GTGACACAAACCAATTTGCAATTTGGGCTGGTTCATATGAGGGTACGTAACCTCAATG 1800
Db 1880 GTGACACAAACCAATTTGCAATTTGGGCTGGTTCATATGAGGGTACGTAACCTCAATG 1939
QY 1801 GTCTCTGGGATCGGGAAGTGGCGTGTTCAAAGTGTGGAATAGCGCTGGCGCTGTATCCCGG 1860
Db 1940 GTCTCTGGGATCGGGAAGTGGCGTGTTCAAAGTGTGGAATAGCGCTGGCGCTGTATCCCGG 1999
QY 1861 TGGGAGTACTATGACTCAGTGTTACAGAACTGTTACATGGGTCTCCCAACTCCAGAAAGAC 1920
Db 2000 TGGGAGTACTATGACTCAGTGTTACAGAACTGTTACATGGGTCTCCCAACTCCAGAAAGAC 2059
QY 1921 AACCTTCACCAATTCAGAAATTTCAACAGTCAATGAGCAGAGCTGAATAATTTTAAACAAGTT 1980
Db 2060 AACCTTCACCAATTCAGAAATTTCAACAGTCAATGAGCAGAGCTGAATAATTTTAAACAAGTT 2119
QY 1981 GAGTACCTCTCTTATTTATGAAACAGCAGATGATAACCTTTCACTTTTCAGCAGTCACTCAG 2040
Db 2120 GAGTACCTCTCTTATTTATGAAACAGCAGATGATAACCTTTCACTTTTCAGCAGTCACTCAG 2179
QY 2041 ATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2100
Db 2180 ATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2239
QY 2101 GACCATGGATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160
Db 2240 GACCATGGATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2299
QY 2161 ATAAAAACAATGTTTCTCTTTACCT 2184
Db 2300 ATAAAAACAATGTTTCTCTTTACCT 2323

RESULT 13
ADJ74790

ID ADJ74790 standard; DNA; 3445 BP.

XX AC ADJ74790;

XX DT 20-MAY-2004 (first entry)

XX DE Marker gene SEQ ID NO:42.

XX KW bronchial asthma; chronic obstructive pulmonary disease;

XX KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;

KW gene therapy; marker gene; gene; ds.

XX Homo sapiens.
 OS
 PN EP1394274-A2.
 XX
 PD 03-MAR-2004.
 XX
 PF 04-AUG-2003; 2003BP-00254857.
 XX
 PR 06-AUG-2002; 2002JP-00229312.
 PR 20-MAR-2003; 2003JP-00077212.
 XX
 XX (GENO-) GENOX RES INC.
 XX
 XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuha K;
 PI WPI; 2004-193155/19.
 DR
 XX
 XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
 PT comparing the expression level of a marker gene in a biological sample
 PT from a subject with the expression level of the gene in a sample from a
 PT healthy subject.
 XX
 XX Claim 1; SEQ ID NO 42; 241bp; English.
 XX
 CC The present invention describes a method of testing for bronchial asthma
 CC or chronic obstructive pulmonary disease. The method comprises
 CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
 CC genes (S1) whose expression levels increase when respiratory epithelial
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
 CC whose expression levels decrease when respiratory epithelial cells are
 CC stimulated with interleukin-13. Also described: (1) a reagent (1) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;
 CC (2) a kit for screening for a candidate compound for a therapeutic agent
 CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
 CC method for producing an animal model for bronchial asthma or chronic
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the
 CC expression of the gene through an RNAi effect or an antibody recognising
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilised to assay a marker gene. (1) has respiratory
 CC and antiasthmatic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for
 CC bronchial asthma or chronic obstructive pulmonary disease. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 XX Sequence 3445 BP; 1085 A; 689 C; 721 G; 950 T; 0 U; 0 Other;
 SQ

Query Match 99.9%; Score 2182.4; DB 12; Length 3445;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 AGTCGAAAACCTTACACTCTAAGTATCTCTACAAACAGAAATATCTTATAGCTGAAGTTATAC 60
 225 AGTCGAAAACCTTACACTCTAAGTATCTCTACAAACAGAAATATCTTATAGCTGAAGTTATAC 284
 61 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAGAAATATCTTGGTGA 120
 285 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAGAAATATCTTGGTGA 344
 121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
 345 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 404

QY 181 GGACATCTCTCAATGATTATTCAATATCTCTGATGGCAGTTTATTCTCTTAGAATAC 240
 Db 405 GGACATCTCTCAATGATTATTCAATATCTCTGATGGCAGTTTATTCTCTTAGAATAC 464
 QY 241 AACTACGTGAAGCAATGGAGCATTCCTTACACAGCTTCATATGACATTTATGATTTAAAT 300
 Db 465 AACTACGTGAAGCAATGGAGCATTCCTTACACAGCTTCATATGACATTTATGATTTAAAT 524
 QY 301 AAAGGACGCTGATTACAGAGAGAGGATTCCTAACACACACACAGTGGTGCATGGTCA 360
 Db 525 AAAGGACGCTGATTACAGAGAGAGGATTCCTAACACACACACAGTGGTGCATGGTCA 584
 QY 361 CCAGTGGGTCTAAATTCGCATATGTTTGGAAACATGACATTTATGTTAAATTCGAACCA 420
 Db 585 CCAGTGGGTCTAAATTCGCATATGTTTGGAAACATGACATTTATGTTAAATTCGAACCA 644
 QY 421 AATTATCCAAAGTTACAGAAATCACTGGAACGGGGAAGAGATATATATATGGAATA 480
 Db 645 AATTATCCAAAGTTACAGAAATCACTGGAACGGGGAAGAGATATATATATGGAATA 704
 QY 481 ACTGACTGGGTATTGAGAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGTTGGTGGTCTCCA 540
 Db 705 ACTGACTGGGTATTGAGAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGTTGGTGGTCTCCA 764
 QY 541 AACGGCACTTTTTTAGCATATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTGAATAC 600
 Db 765 AACGGCACTTTTTTAGCATATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTGAATAC 824
 QY 601 TCCTTTCTACTCTGATGAGTCACTGCGATACCCAAAGACTGTACGGGTTCCTATATCCAAAG 660
 Db 825 TCCTTTCTACTCTGATGAGTCACTGCGATACCCAAAGACTGTACGGGTTCCTATATCCAAAG 884
 QY 661 CGAGAGCTGTGAATCCAACTGTGAAGTCTTTGTTGTTAAATACAGACTCTCTCAGCTCA 720
 Db 885 CGAGAGCTGTGAATCCAACTGTGAAGTCTTTGTTGTTAAATACAGACTCTCTCAGCTCA 944
 QY 721 GTCCCAATGCAACTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGATCAC 780
 Db 945 GTCCCAATGCAACTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGATCAC 1004
 QY 781 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGTCTCAGGAGG 840
 Db 1005 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGTCTCAGGAGG 1064
 QY 841 ATTGAAATCTATTGGTTCATGATATTTGTCATGATGATCCAGTGGAGATGGAC 900
 Db 1065 ATTGAAATCTATTGGTTCATGATATTTGTCATGATGATCCAGTGGAGATGGAAAC 1124
 QY 901 TGCTTAGTGGCACGGCAACACATTCGAAATGAGTACTTCTGCTGGGTGGGAGATTTAGG 960
 Db 1125 TGCTTAGTGGCACGGCAACACATTCGAAATGAGTACTTCTGCTGGGTGGGAGATTTAGG 1184
 QY 961 CCTTCAGAACCTCAATTTTACCCTTGTGATGATTAAGCTTCTCAAGATCATCAGCAATGAA 1020
 Db 1185 CCTTCAGAACCTCAATTTTACCCTTGTGATGATTAAGCTTCTCAAGATCATCAGCAATGAA 1244
 QY 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAGACTGCACATTTATTACA 1080
 Db 1245 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAGACTGCACATTTATTACA 1304
 QY 1081 AAAGGCACTGGGAAGTCACTGGGATAGAGCTCTAACAGTGAATATCTATCTACTACATTT 1140
 Db 1305 AAAGGCACTGGGAAGTCACTGGGATAGAGCTCTAACAGTGAATATCTATCTACTACATTT 1364
 QY 1141 AGTATGAATATAAAGGAATCCAGGAGGAAGGAATCTTTATATAAATCCAACTTTATTGAC 1200
 Db 1365 AGTATGAATATAAAGGAATCCAGGAGGAAGGAATCTTTATATAAATCCAACTTTATTGAC 1424
 QY 1201 TATACAAAGTGAATGCTCAGTGTGAGCTGAATCCGGAAAGGTGTGCTAGTACTATTCT 1260
 Db 1425 TATACAAAGTGAATGCTCAGTGTGAGCTGAATCCGGAAAGGTGTGCTAGTACTATTCT 1484
 QY 1261 GTGTCAATTCAGTAAAGAGCGCAAGTATTATCAGCTGAGATGTTCCGGTCTCTGCTGCC 1320


```
Db 1485 GTGTCTATTTCAGTAAAGGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTCGGTCTGCC 1544
Qy 1321 CTCCTATCTCTACACAGCAGCGGTGAATGATGAATAAGGCTGAGAGTCTCTGGAAGACAATTC 1380
Db 1545 CTCCTATCTCTACACAGCAGCGGTGAATGATGAATAAGGCTGAGAGTCTCTGGAAGACAATTC 1604
Qy 1381 GCTTTGGATAAATGCTCGAGATGTCAGATGTCAGATGTCCTCCAAAACCTGGACTTCATTATT 1440
Db 1605 GCTTTGGATAAATGCTCGAGATGTCAGATGTCCTCCAAAACCTGGACTTCATTATT 1664
Qy 1441 TTGAATGAACAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGGATAAATCCAAAG 1500
Db 1665 TTGAATGAACAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGGATAAATCCAAAG 1724
Qy 1501 AAATATCTCTACTATTAGATGTTATGTCAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1560
Db 1725 AAATATCTCTACTATTAGATGTTATGTCAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1784
Qy 1561 TTCAGACTGAACCTGGCCACTTACCTTGCAGACAGAAAACATTTATAGTAGTACTGCTTT 1620
Db 1785 TTCAGACTGAACCTGGCCACTTACCTTGCAGACAGAAAACATTTATAGTAGTACTGCTTT 1844
Qy 1621 GATGCGCAGAGAGTGTTCACAGGAGATAGATCATGATGCAATCAACAGAGACTG 1680
Db 1845 GATGCGCAGAGAGTGTTCACAGGAGATAGATCATGATGCAATCAACAGAGACTG 1904
Qy 1681 GGAACATTTGAAGTTGAAGATCAAAATGAAGCAGCCAGACAATTTCAAAAATGGATTT 1740
Db 1905 GGAACATTTGAAGTTGAAGATCAAAATGAAGCAGCCAGACAATTTCAAAAATGGATTT 1964
Qy 1741 GTGGACAACAAACGAATTTGGGGCTGGTCAATGATGAGGGTACGTAACTCAATG 1800
Db 1965 GTGGACAACAAACGAATTTGGGGCTGGTCAATGATGAGGGTACGTAACTCAATG 2024
Qy 1801 GTCCTGGGATCGGAAGTGGCGGTTCAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1860
Db 2025 GTCCTGGGATCGGAAGTGGCGGTTCAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 2084
Qy 1861 TGGGAGTACTATGACTCAGTGTCACAGAACTTACATGGGTCTCCCACTCCAGAGAC 1920
Db 2085 TGGGAGTACTATGACTCAGTGTCACAGAACTTACATGGGTCTCCCACTCCAGAGAC 2144
Qy 1921 AACCTTGACCAATTACAGAAATTCACAGTCAATGACGAGCTGAAAATTTTAAACAAGTT 1980
Db 2145 AACCTTGACCAATTACAGAAATTCACAGTCAATGACGAGCTGAAAATTTTAAACAAGTT 2204
Qy 1981 GAGTACTCTTTATTCATGGAACAGCAGATGATAACTTCACTTTCAGCAGTCAAGTCAAG 2040
Db 2205 GAGTACTCTTTATTCATGGAACAGCAGATGATAACTTCACTTTCAGCAGTCAAGTCAAG 2264
Qy 2041 ATCTCCAAAGCCCTGGTGGATGTTGGAGTGGATTTCCAGGCAATGTTGTTATGATGAA 2100
Db 2265 ATCTCCAAAGCCCTGGTGGATGTTGGAGTGGATTTCCAGGCAATGTTGTTATGATGAA 2324
Qy 2101 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160
Db 2325 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2384
Qy 2161 ATAAAACAATGTTTCTCTTTTACCT 2184
Db 2385 ATAAAACAATGTTTCTCTTTTACCT 2408
```

RESULT 14

ADO19397

ID ADO19397 standard; cDNA; 3445 BP.

XX

AC ADO19397;

XX

DT 12-AUG-2004 (first entry)

XX

DE Human PRO polynucleotide #164.

```
XX Human; PRO; gene; ss; immune related disorder;
KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
KW renal disease; demyelinating disease; central nervous system;
KW peripheral nervous system; demyelinating polyneuropathy;
KW Guillain-Barre syndrome;
XX Chronic inflammatory demyelinating polyneuropathy.
XX Homo sapiens.
OS WO2004043361-A2.
XX 27-MAY-2004.
XX 06-NOV-2003; 2003WO-US035268.
XX 08-NOV-2002; 2002US-0425235P.
XX (GETH ) GENENTECH INC.
XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX WPI: 2004-420067/39.
XX P-PSDB; ADOI9398.
XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthritis.
XX Claim 1; SEQ ID NO 340; 1731pp; English.
XX The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The polypeptides and polynucleotides are useful for
CC treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the
CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polynucleotide of
CC the invention.
XX SQ Sequence 3445 BP; 1085 A; 689 C; 721 G; 950 T; 0 U; 0 Other;
```

Query Match 99.9%; Score 2182.4; DB 12; Length 3445;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 AGTCGCAAACTTACACTCTAACTGATTAATAAATACTTATAGACTGAAGTTATAC 60
Db 225 AGTCGCAAACTTACACTCTAACTGATTAATAAATACTTATAGACTGAAGTTATAC 284
Qy 61 TCCTTAAGATGGATTTTCAGATCATGATATCTCTACAAACAAGAAAAATAATATCTTGGTA 120
Db 285 TCCTTAAGATGGATTTTCAGATCATGATATCTCTACAAACAAGAAAAATAATATCTTGGTA 344
Qy 121 TTCATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
Db 345 TTCATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 404
Qy 181 GGACATTTCTATCAATGATTAATCAATATCTCCCTGATGGCGAGTTTATCTCTTAGAATAC 240
Db 405 GGACATTTCTATCAATGATTAATCAATATCTCCCTGATGGCGAGTTTATCTCTTAGAATAC 464
Qy 241 AACTTACGTGAAGCAATGGAGGCATTCCTACACAGCTTCATATGACATTTATGATTAAAT 300
```

Db 465 AACTACGTGAAGCAATGGAGGCATTCCTACACAGCTTCATATGACATTTATGATTAAAT 524
Qy 301 AAAAGCAGCTGATTACAGAGAGAGGATTCCAAACAAACACACACAGTGGGTCACTGGTCA 360
Db 525 AAAAGCAGCTGATTACAGAGAGAGGATTCCAAACAAACACACACAGTGGGTCACTGGTCA 584
Qy 361 CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCA 420
Db 585 CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCA 644
Qy 421 AATTATACCAAGTTTACAGAAATCACATGGACGGGAAAGAGATATATATATATGAAATA 480
Db 645 AATTATACCAAGTTTACAGAAATCACATGGACGGGAAAGAGATATATATATATGAAATA 704
Qy 481 ACTGACTGGGTTTATGAAGAGAGGTCCTTCACTGCTCTCTGCTCTGCTGCTGCTGCTGCT 540
Db 705 ACTGACTGGGTTTATGAAGAGAGGTCCTTCACTGCTCTCTGCTCTGCTGCTGCTGCTGCT 764
Qy 541 AACGGCACTTTTATGACATATGCCCAATTTTAAACGACACAGAGTCCCACTTATGAAATAC 600
Db 765 AACGGCACTTTTATGACATATGCCCAATTTTAAACGACACAGAGTCCCACTTATGAAATAC 824
Qy 601 TCCTTCTACTCTGATGAGTCACCTGACAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 660
Db 825 TCCTTCTACTCTGATGAGTCACCTGACAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 884
Qy 661 GCAGGAGCTGTAATCCAACTGTAAAGTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720
Db 885 GCAGGAGCTGTAATCCAACTGTAAAGTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 944
Qy 721 GTACCAATGCAACTTCCATACAAATCACTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 780
Db 945 GTACCAATGCAACTTCCATACAAATCACTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1004
Qy 781 TACTTGTGTGATGTGATGAGTAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
Db 1005 TACTTGTGTGATGTGATGAGTAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1064
Qy 841 ATTACAGAACTATTTCGGTCAATGATATTTGTGATATGATGATGATGATGATGATGATGATGAT 900
Db 1065 ATTACAGAACTATTTCGGTCAATGATATTTGTGATATGATGATGATGATGATGATGATGATGAT 1124
Qy 901 TCGTTAGTGGCAGCGCAACATTTGAAATGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 1125 TCGTTAGTGGCAGCGCAACATTTGAAATGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1184
Qy 961 CCTTCAGAACCTCATTTTACCTTGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1020
Db 1185 CCTTCAGAACCTCATTTTACCTTGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1244
Qy 1021 GAAGGTTACAGACACATTTTGTCTATTTTCCAAATAGATAAAAGACTGCACTTTATTATTA 1080
Db 1245 GAAGGTTACAGACACATTTTGTCTATTTTCCAAATAGATAAAAGACTGCACTTTATTATTA 1304
Qy 1081 AAAGGCACCTGGGAAGTCACTGGGATAGAGCTCTAACAGTGATTTATCTATCTACATTT 1140
Db 1305 AAAGGCACCTGGGAAGTCACTGGGATAGAGCTCTAACAGTGATTTATCTATCTACATTT 1364
Qy 1141 AGTAATGAATATATAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAAAATCCAACTTATTGAC 1200
Db 1365 AGTAATGAATATATAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAAAATCCAACTTATTGAC 1424
Qy 1201 TATACAAAAGTGACATGCTCCTGATGAGTGAATCCCGAAAGGTTGTCAGTACTATTCTT 1260
Db 1425 TATACAAAAGTGACATGCTCCTGATGAGTGAATCCCGAAAGGTTGTCAGTACTATTCTT 1484
Qy 1261 GTGTCAATTCAGTAAAGGCGAGTATATCACTGCTGAGATGTTCCCGTCTGCTGCTGCC 1320
Db 1485 GTGTCAATTCAGTAAAGGCGAGTATATCACTGCTGAGATGTTCCCGTCTGCTGCTGCC 1544
Qy 1321 CTCTATCTCTACACAGCAGCTGAATGATTAAGGGCTGAGAGTCTCTGGAAGCAATTTCA 1380
Db 1545 CTCTATCTCTACACAGCAGCTGAATGATTAAGGGCTGAGAGTCTCTGGAAGCAATTTCA 1604

Qy 1381 GCTTTGGATATAAATGCTCCAGAAATGCTCCAGATGCCCTCCAAAAAATCGACTTCAATTAT 1440
Db 1605 GCTTTGGATATAAATGCTCCAGAAATGCTCCAGATGCCCTCCAAAAAATCGACTTCAATTAT 1664
Qy 1441 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTTGGCTCTCTCAATTTTGTATAAATCCAAG 1500
Db 1665 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTTGGCTCTCTCAATTTTGTATAAATCCAAG 1724
Qy 1501 AAATATCTCTACTATTAATAGATGTTGAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1560
Db 1725 AAATATCTCTACTATTAATAGATGTTGAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1784
Qy 1561 TTCAGACTGAACTGGGCCACTTACCTTCCAGAGCAGAGAAAACATTTATAGTACTAGCTTT 1620
Db 1785 TTCAGACTGAACTGGGCCACTTACCTTCCAGAGCAGAGAAAACATTTATAGTACTAGCTTT 1844
Qy 1621 GATGCGCAGAGGAAGTGGTTTACCAAGGAGATGAAGATCATGCAATCAACAGAGAGACTG 1680
Db 1845 GATGCGCAGAGGAAGTGGTTTACCAAGGAGATGAAGATCATGCAATCAACAGAGAGACTG 1904
Qy 1681 GGAAATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1740
Db 1905 GGAAATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1964
Qy 1741 GTGGACAAACAAAGAAATTTGGGCTGGTTCATATGAGGGGTACGTAACTCAATG 1800
Db 1965 GTGGACAAACAAAGAAATTTGGGCTGGTTCATATGAGGGGTACGTAACTCAATG 2024
Qy 1801 GTCTCTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1860
Db 2025 GTCTCTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 2084
Qy 1861 TGGGAGTACTATGACTCAGTGTACAGAACTGTTACATGGGTCTCCCACTCCAGAGAC 1920
Db 2085 TGGGAGTACTATGACTCAGTGTACAGAACTGTTACATGGGTCTCCCACTCCAGAGAC 2144
Qy 1921 AACCTTGACCAATTACAGAAATTTCAACAGTCAAGCAGAGCTGAAAAATTTTAAAAACAAGTT 1980
Db 2145 AACCTTGACCAATTACAGAAATTTCAACAGTCAAGCAGAGCTGAAAAATTTTAAAAACAAGTT 2204
Qy 1981 GAGTACCTCTTATTTATGGAACAGCAGATGATAACGTTTCACTTTTCCAGCAGTCACTCAG 2040
Db 2205 GAGTACCTCTTATTTATGGAACAGCAGATGATAACGTTTCACTTTTCCAGCAGTCACTCAG 2264
Qy 2041 ATCTCCAAAGCCCTGGTCAATGTTGGAGTGGATTTTCCAGGCAATGTGGTATGATGAA 2100
Db 2265 ATCTCCAAAGCCCTGGTCAATGTTGGAGTGGATTTTCCAGGCAATGTGGTATGATGAA 2324
Qy 2101 GACCATGGAATAGCTAGCAGCAGCAGACCAACATATATATACCCATGAGCCACTTTC 2160
Db 2325 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2384
Qy 2161 ATAAAAACAATGTTTCTCTTTTACCT 2184
Db 2385 ATAAAAACAATGTTTCTCTTTTACCT 2408

RESULT 15

ADO71613

ID ADO71613 standard; DNA; 3445 BP.

XX ADO71613;

XX ADO71613;

DT 26-AUG-2004 (first entry)

XX Nucleotide sequence of a human CD26 protein.

XX CD26; chemotherapeutic; radiotherapeutic; cancer; cell growth;

XX dipeptidyl peptidase IV; DPPIV; topoisomerase II inhibitor;

KW cell cycle arrest; tumour; tumour necrosis; immune response; human; gene;

XX SB.

OS Homo sapiens:
 FH Key Location/Qualifiers
 FT CDS 111..2411
 FT /*tag= a
 FT /product= "CD26"
 FN WO2004045497-A2.
 PD 03-JUN-2004.
 PF 15-MAY-2003; 2003WO-US015499.
 XX 17-MAY-2002; 2002US-0381606P.
 PR (TEXA) UNIV TEXAS SYSTEM.
 PA Dang NH, Morimoto C;
 PI WPI; 2004-420511/39.
 DR P-PSDB; ADO71612.
 DR
 DR
 PT Use of a CD26 composition, and a chemotherapeutic and/or a
 PT radiotherapeutic agent for e.g. inhibiting the cell growth, inducing cell
 PT cycle arrest, killing a cancer cell, treating cancer, or inducing tumor
 PT regression or tumor necrosis.
 XX
 XX Claim 14; Page 153-157; 182pp; English.
 PS
 CC The specification describes a CD26 composition which, in conjunction with
 CC chemotherapeutic or radiotherapeutic agents, is used for the treatment
 CC and prevention of cancers. Expression of CD26 enhances the sensitivity of
 CC the cancer cell to the chemotherapeutic or radiotherapeutic agent. CD26
 CC is a dipeptidyl peptidase IV (DPP-IV). The chemotherapeutic agent is a
 CC topoisomerase II inhibitor. The CD26 composition of the invention is
 CC useful for inhibiting the growth of a cell, inducing cell cycle arrest in
 CC a cell, killing a cancer cell, potentiating the effect of a
 CC chemotherapeutic agent and/or a radiotherapeutic agent on a tumor cell,
 CC inducing or enhancing apoptosis of a cancer cell, treating cancer, or
 CC further useful for increasing topoisomerase II expression in a cell, for
 CC activating an antigen-presenting cell, or for potentiating immune
 CC responses of an animal. The present sequence encodes a CD26 protein, and
 CC is used in vectors which are used to produce compositions of the
 CC invention.
 XX
 SQ Sequence 3445 BP; 1085 A; 689 C; 721 G; 950 T; 0 U; 0 Other;
 Query Match 99.9%; Score 2182.4; DB 12; Length 3445;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AGTCGCAAACTTACACTCTTACCTGATTAATTAATAAATACTTATAGCTGAAGTTATAC 60
 DB 225 AGTCGCAAACTTACACTCTTACCTGATTAATTAATAAATACTTATAGCTGAAGTTATAC 284
 QY 61 TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAACAAGAAATAATATCTTGTA 120
 DB 285 TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAACAAGAAATAATATCTTGTA 344
 QY 121 TTCATGCTGAATATGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
 DB 345 TTCATGCTGAATATGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT 404
 QY 181 GGACATTTCTATCAATGATTAATCAATATCTCTGATGGGAGTTTATCTCTTAGAATAC 240
 DB 405 GGACATTTCTATCAATGATTAATCAATATCTCTGATGGGAGTTTATCTCTTAGAATAC 464
 QY 241 AACTACGTTGAAGCAATGAGGCAATTCCTACACGCTTCATATGACATTTATGATTAAT 300
 DB 465 AACTACGTTGAAGCAATGAGGCAATTCCTACACGCTTCATATGACATTTATGATTAAT 524
 QY 301 AAAAGGCAGCTGATTACAGAGAGAGGATTCCAAACAACACACAGTGGGTCACTAGTCA 360

DB 525 AAAAGGCAGCTGATTACAGAGAGAGGATTCCAAACAACACACAGTGGGTCACTAGTCA 584
 QY 361 CCAGTGGGTCAATAAATTGGCATATGTTTGGACACATGACATTTATGTTAAATTTGAACCA 420
 DB 585 CCAGTGGGTCAATAAATTGGCATATGTTTGGACACATGACATTTATGTTAAATTTGAACCA 644
 QY 421 AATTTACCAAGTTACAGAAATCAATGACGGGGAAGAAAGATATAATATAATGAATA 480
 DB 645 AATTTACCAAGTTACAGAAATCAATGACGGGGAAGAAAGATATAATATAATGAATA 704
 QY 481 ACTGACTGGGTTTATGAAGAGAAAGTCTTCACTGCTCTCTGCTCTGCTGCTCTCA 540
 DB 705 ACTGACTGGGTTTATGAAGAGAAAGTCTTCACTGCTCTCTGCTCTGCTGCTCTCA 764
 QY 541 AACGGCACTTTTATGAGCATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATGAATAC 600
 DB 765 AACGGCACTTTTATGAGCATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATGAATAC 824
 QY 601 TCCTTTCTACTCTGATGAGTCACTGACGATCCCAAGAGCTGTACGGGTTCCATATCCAAAG 660
 DB 825 TCCTTTCTACTCTGATGAGTCACTGACGATCCCAAGAGCTGTACGGGTTCCATATCCAAAG 884
 QY 661 GCAGGAGCTGGAATCCCAACTGTAAAGTTCTTTGTGTPAATAACAGACTCTCTCAGCTCA 720
 DB 885 GCAGGAGCTGGAATCCCAACTGTAAAGTTCTTTGTGTPAATAACAGACTCTCTCAGCTCA 944
 QY 721 GTCAACCAATGCACTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGGATCAC 780
 DB 945 GTCAACCAATGCACTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGGATCAC 1004
 QY 781 TACTTGTGTGATGTGACATGGGCAACACAAAGAAAGATTTCTTTCAGTGGCTCAGGAGG 840
 DB 1005 TACTTGTGTGATGTGACATGGGCAACACAAAGAAAGATTTCTTTCAGTGGCTCAGGAGG 1064
 QY 841 ATTCAAGAACTATTTCGGTCAATGATATTTGTGACTATGATGAATCCAGTGGAAAGATGGAAC 900
 DB 1065 ATTCAAGAACTATTTCGGTCAATGATATTTGTGACTATGATGAATCCAGTGGAAAGATGGAAC 1124
 QY 901 TGCTTACTGCGACCGGCAACATTTGAATGAGTACTACTGCTGGTGGTGGAGATTTAGG 960
 DB 1125 TGCTTACTGCGACCGGCAACATTTGAATGAGTACTACTGCTGGTGGTGGAGATTTAGG 1184
 QY 961 CCTTCAGAACTCAATTTTACCTTTGATGGTAATAGCTTTCTACAAGATCATCAGCAATCAA 1020
 DB 1185 CCTTCAGAACTCAATTTTACCTTTGATGGTAATAGCTTTCTACAAGATCATCAGCAATCAA 1244
 QY 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAGAGCTGACATTTATTTACA 1080
 DB 1245 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAGAGCTGACATTTATTTACA 1304
 QY 1081 AAAGGCACCTGGGAAGTCACTGGGATAGAGCTCTTAACCGATGATTATCTATCTACATTT 1140
 DB 1305 AAAGGCACCTGGGAAGTCACTGGGATAGAGCTCTTAACCGATGATTATCTATCTACATTT 1364
 QY 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCCAACTTATTGAC 1200
 DB 1365 AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCCAACTTATTGAC 1424
 QY 1201 TATACAAAAGTGAACATCCCTCAGTTGTGAGCTGAATCCGGAAGAGGTGTCAGTACTTTCT 1260
 DB 1425 TATACAAAAGTGAACATCCCTCAGTTGTGAGCTGAATCCGGAAGAGGTGTCAGTACTTTCT 1484
 QY 1261 GTGTCATTCAGTAAGAGGCGAAGTATTATCAGCTGAGATGTTCCGGCTCCTGGTCTGCC 1320
 DB 1485 GTGTCATTCAGTAAGAGGCGAAGTATTATCAGCTGAGATGTTCCGGCTCCTGGTCTGCC 1544
 QY 1321 CTCTATCTCTACACAGCAGCTGAATGATAAAGGGCTGAGAGTCTCGGGAAGACAAATTTCA 1380
 DB 1545 CTCTATCTCTACACAGCAGCTGAATGATAAAGGGCTGAGAGTCTCGGGAAGACAAATTTCA 1604
 QY 1381 GCTTTGGATAAAATGCTGCAAGATGTCCAGATGCCCTCCAAAAAATCTGAGCTTCAATTAT 1440

[illegible]

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2006, 04:24:53 ; Search time 11140 Seconds
(without alignments)
11144.174 Million cell updates/sec

Title: US-10-659-055-2
Perfect score: 2184
Sequence: 1 agtcgcaaacattacactct.....acaatgtttctttacct 2184

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_sts.*
- 11: gb_sy.*
- 12: gb_un.*
- 13: gb_vi.*
- 14: gb_htg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2184	100.0	2211	6	CQ815921 Sequence
2	2184	100.0	3407	6	CQ963950 Sequence
3	2184	100.0	3407	6	AR349829 Sequence
4	2184	100.0	3407	6	AR653630 Sequence
5	2184	100.0	3407	8	HSFCHDP7
6	2182.4	99.9	2301	6	CQ772397 Sequence
7	2182.4	99.9	2431	8	BC013329 Homo sapi
8	2182.4	99.9	3411	6	CQ727953 Sequence
9	2182.4	99.9	3445	6	CQ776356 Sequence
10	2182.4	99.9	3445	6	CQ981570 Sequence
11	2182.4	99.9	3445	6	CS031460 Sequence
12	2182.4	99.9	3445	6	CS032879 Sequence
13	2182.4	99.9	3445	6	CS040412 Sequence
14	2182.4	99.9	3445	6	CS041831 Sequence
15	2182.4	99.9	3913	8	BC065265 Homo sapi
16	2182.4	99.9	3948	6	CS134667 Sequence
17	2180.8	99.9	2924	6	AR380481 Sequence
18	2180.8	99.9	2924	8	HUMCD26A M74777 Human dipep

19	2177.6	99.7	2375	8	HUMDIPBPV
20	1852.8	84.8	2573	4	AB023952
21	1844.8	84.5	2727	4	AF461806
22	1843.2	84.4	2298	4	AY056834
23	1827.8	83.7	2400	4	AY198323
24	1677.6	76.8	3316	6	CQ777283
25	1677.6	76.8	3316	9	MMDDPP4
26	1677.6	76.8	3482	9	BC022183
27	1657.2	75.9	3214	9	RATGPMBCD
28	1656.2	75.8	4835	6	AX401894
29	1656.2	75.8	4835	9	RATDPP
30	1009.2	46.2	3203	5	AJ720903
31	923.2	42.3	4250	5	AB158225
32	923.2	42.3	4254	5	AB158224
33	762.4	34.9	2834	5	CR926279
34	761.6	34.9	3376	5	BC082401
35	754.2	34.5	2366	8	HSU76833
36	752.6	34.5	2648	8	BC026250
37	751	34.4	2814	6	AX329990
38	751	34.4	2814	6	AX330925
39	751	34.4	2814	6	AX333666
40	751	34.4	2814	6	AX335485
41	751	34.4	2814	6	AX578046
42	751	34.4	2814	8	HSU09278
43	751	34.4	2815	6	AR078938
44	751	34.4	2815	6	I32308
45	751	34.4	2815	6	AR633310

ALIGNMENTS

RESULT 1
LOCUS CQ815921 2211 bp DNA linear PAT 03-JUN-2004
DEFINITION Sequence 1 from Patent EPI422293.
ACCESSION CQ815921
VERSION CQ815921.1 GI:48144442
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Hennig, M., Loeffler, B. M. and Thoma, R.
TITLE Crystal structure of dipeptidyl peptidase IV
JOURNAL Patent: EP 1422293-A 1 26-MAY-2004;
F. HOFFMANN-LA ROCHE AG (CH)
FEATURES
source Location/Qualifiers
1..2211 /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 2184; DB 6; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AGTCGCAAACTTACACTCTAACTGATTA	CTTAAAAAATCTTATAGACGAACTTATAC	60
Db	25	AGTCGCAAACTTACACTCTAACTGATTA	CTTAAAAAATCTTATAGACGAACTTATAC	84
Qy	61	TCCTTAAGATGGATTTTCAGATCATGAT	ATCTCTACAAACAGAAATATATCTTGGTA	120
Db	85	TCCTTAAGATGGATTTTCAGATCATGAT	ATCTCTACAAACAGAAATATATCTTGGTA	144
Qy	121	TTCAATGCTCAATATGGAACAGCTAGT	TTTCTTGGAGAACAGTACATTTGAGAGTTT	180
Db	145	TTCAATGCTCAATATGGAACAGCTAGT	TTTCTTGGAGAACAGTACATTTGAGAGTTT	204
Qy	181	GGACATTTCTATCAATGATTTTCAATAT	CTCTCTGATGGCGAGTTTATTTCTTCTT	240

Db 205 GGCATCTTCATGATTTTCAATATCTCTGATGGCAGTTATTTCTTTAGATAC 264
 Qy 241 AACTACGTGAAGCAATGGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300
 Db 265 AACTACGTGAAGCAATGGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 324
 Qy 301 AAAAGGCAGCTGATTTACAGAGAGAGGATTCACAAACACACACAGTGGGTCAATGGTCA 360
 Db 325 AAAAGGCAGCTGATTTACAGAGAGAGGATTCACAAACACACACAGTGGGTCAATGGTCA 384
 Qy 361 CAGTGGGTCAATAATGGCATATGTTGGAAACAATGACATTTATGTTAAAAATGAAACA 420
 Db 385 CAGTGGGTCAATAATGGCATATGTTGGAAACAATGACATTTATGTTAAAAATGAAACA 444
 Qy 421 AATTTTACCAAGTTTACAGATATCATGGACGGGGAAGAGATATATATATATATGAAATA 480
 Db 445 AATTTTACCAAGTTTACAGATATCATGGACGGGGAAGAGATATATATATATATGAAATA 504
 Qy 481 ACTGACTGGGTTTATGAGAGAGAGTCTTCAGTGGCTACTCTGCTCTGTTGGTGGTCTCA 540
 Db 505 ACTGACTGGGTTTATGAGAGAGAGTCTTCAGTGGCTACTCTGCTCTGTTGGTGGTCTCA 564
 Qy 541 AACGGCACTTTTATGACATATGCCCAATTTAACGACACACAGAGTCCCACTTATGAAATAC 600
 Db 565 AACGGCACTTTTATGACATATGCCCAATTTAACGACACACAGAGTCCCACTTATGAAATAC 624
 Qy 601 TCCTTCTACTCTGATGAGTCACTGACGATACCAAGACATGTAAGGGTTCATATCAAAG 660
 Db 625 TCCTTCTACTCTGATGAGTCACTGACGATACCAAGAGCTGTACGGGTTCATATCAAAG 684
 Qy 661 GCAGGAGCTGTGAATCCAACTGTAAGTCTTTGTTGTAATACACAGCTCTCTAGCTCA 720
 Db 685 GCAGGAGCTGTGAATCCAACTGTAAGTCTTTGTTGTAATACACAGCTCTCTAGCTCA 744
 Qy 721 GTCCCAATGCAACTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGATCAC 780
 Db 745 GTCCCAATGCAACTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGATCAC 804
 Qy 781 TACTTGTGTGATGTGATGAGGCAACACAAAGAAAGAAATTTCTTTCAGTGGTCTAGGAG 840
 Db 805 TACTTGTGTGATGTGATGAGGCAACACAAAGAAAGAAATTTCTTTCAGTGGTCTAGGAG 864
 Qy 841 ATTCCAGAACTATTCGGTCAATGATATTTGTGACTATGATGATCCAGTGGAGATGGAC 900
 Db 865 ATTCCAGAACTATTCGGTCAATGATATTTGTGACTATGATGATCCAGTGGAGATGGAC 924
 Qy 901 TGCCTTAGTGGCACGGCAACACATTTGAAATGAGTACTACTGCTGGGTGGGTTGGAAGATTTAGG 960
 Db 925 TGCCTTAGTGGCACGGCAACACATTTGAAATGAGTACTACTGCTGGGTGGGTTGGAAGATTTAGG 984
 Qy 961 CTTTCAGAACTCATTTTACCTTACCTTGAATGATGCTTCTCAAGATCATCAGCAATGAA 1020
 Db 985 CTTTCAGAACTCATTTTACCTTACCTTGAATGATGCTTCTCAAGATCATCAGCAATGAA 1044
 Qy 1021 GAAGGTTACAGACATTTGCTATTTCCAAATAGATATAAAAGACCTGACATTTATTTACA 1080
 Db 1045 GAAGGTTACAGACATTTGCTATTTCCAAATAGATATAAAAGACCTGACATTTATTTACA 1104
 Qy 1081 AAAGGCACCTGGGAGTCAATCGGATAGAGTCTTAAACAGTGAATTTATCTATCTACATTT 1140
 Db 1105 AAAGGCACCTGGGAGTCAATCGGATAGAGTCTTAAACAGTGAATTTATCTATCTACATTT 1164
 Qy 1141 AGTAATGAATATAAAGGAATCCAGGAGGAAGAAATCTTTTAAATAATCAATTTATGAC 1200
 Db 1165 AGTAATGAATATAAAGGAATCCAGGAGGAAGAAATCTTTTAAATAATCAATTTATGAC 1224
 Qy 1201 TATACAAAGTGAATGCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1260
 Db 1225 TATACAAAGTGAATGCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1284
 Qy 1261 GTGTCAATTCAGTAAAGAGGCGAGTATTTATCAGCTGAGATGTTCCGGTCTCTGCTGCCCC 1320

Db 1285 GTGTCAATTCAGTAAAGAGGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTGCTGCCCC 1344
 Qy 1321 CTCTATATCTTACACAGCAGCGTGAATGATATAAAGGCTGAGAGTCTCTGGAAGACAATTTCA 1380
 Db 1345 CTCTATATCTTACACAGCAGCGTGAATGATATAAAGGCTGAGAGTCTCTGGAAGACAATTTCA 1404
 Qy 1381 GCTTTGGATAAAATGCTCCAGAAATGTCAGATGTCCTCCAAAAAACTGGACTTTCAATTT 1440
 Db 1405 GCTTTGGATAAAATGCTCCAGAAATGTCAGATGTCCTCCAAAAAACTGGACTTTCAATTT 1464
 Qy 1441 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCAATTTTGAATAAATCCAAAG 1500
 Db 1465 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCAATTTTGAATAAATCCAAAG 1524
 Qy 1501 AAATATCTCTACTATTTAGATGTTGATGCGAGGCCCATGCTAGTCAAAAAGCAGACACTGTC 1560
 Db 1525 AAATATCTCTACTATTTAGATGTTGATGCGAGGCCCATGCTAGTCAAAAAGCAGACACTGTC 1584
 Qy 1561 TTGACACTGAACTGGGCCACTTTACCTTGCAAAGCACAGAAACATTTATAGTAGTACTAGCTTT 1620
 Db 1585 TTGACACTGAACTGGGCCACTTTACCTTGCAAAGCACAGAAACATTTATAGTAGTACTAGCTTT 1644
 Qy 1621 GATGCGACAGAAAGTGGTTTACAAAGAGATGAAGATCATGATGCAATCAACAGAGACTG 1680
 Db 1645 GATGCGACAGAAAGTGGTTTACAAAGAGATGAAGATCATGATGCAATCAACAGAGACTG 1704
 Qy 1681 GGAACATTTGAGTTGAGATCMAATTTGAAGCAGCAGACCAATTTTCAAAAATGGATTT 1740
 Db 1705 GGAACATTTGAGTTGAGATCMAATTTGAAGCAGCAGACCAATTTTCAAAAATGGATTT 1764
 Qy 1741 GTGAGACAAACAAACGAATTTGGGGCTGGTCAATGAGAGGATGCTAACTCAATG 1800
 Db 1765 GTGAGACAAACAAACGAATTTGGGGCTGGTCAATGAGAGGATGCTAACTCAATG 1824
 Qy 1801 GTCTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1860
 Db 1825 GTCTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1884
 Qy 1861 TGGGATGACTATGACTGATGATGACAGAACTTTCAACAGTCAATGAGAGGATGCTAACTCAATG 1920
 Db 1885 TGGGATGACTATGACTGATGATGACAGAACTTTCAACAGTCAATGAGAGGATGCTAACTCAATG 1944
 Qy 1921 AACCTTGACCAATTTACAGAAATTTCAACAGTCAATGAGAGGATGCTAACTCAATG 1980
 Db 1945 AACCTTGACCAATTTACAGAAATTTCAACAGTCAATGAGAGGATGCTAACTCAATG 2004
 Qy 1981 GAGTACCTCTTATTTATGAGAAACAGACATGATGATGATGATGATGATGATGATGATGATGAT 2040
 Db 2005 GAGTACCTCTTATTTATGAGAAACAGACATGATGATGATGATGATGATGATGATGATGATGAT 2064
 Qy 2041 ATCTCCAAAGCCCTGGTGGATGTTGGAGTGGATTTCCAGGCAATGCTGATGATGATGATGATGAT 2100
 Db 2065 ATCTCCAAAGCCCTGGTGGATGTTGGAGTGGATTTCCAGGCAATGCTGATGATGATGATGATGAT 2124
 Qy 2101 GACCATGGAATAGCTAGCAGCAGACACACAAACATATATATATATATATATATATATATATATAT 2160
 Db 2125 GACCATGGAATAGCTAGCAGCAGACACACAAACATATATATATATATATATATATATATATATAT 2184
 Qy 2161 ATAAACCAATGTTCTCTTTTACCT 2184
 Db 2185 ATAAACCAATGTTCTCTTTTACCT 2208

RESULT 2
 CQ963950
 LOCUS
 DEFINITION
 CQ963950
 ACCESSION
 CQ963950.1
 VERSION
 CQ963950.1
 GI:56562717
 KEYWORDS
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CQ963950
 Sequence 1 from Patent WO2004104216.
 linear PAT 13-DBC-2004

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
Hominidae; Homo.	
1	
REFERENCE	
AUTHORS	Golz,S., Brueggemeier,U. and Summer,H.
TITLE	Diagnostics and therapeutics for diseases associated with
JOURNAL	dipeptidylpeptidase iv (dpp4)
Patent:	WO 2004104216-A 1 02-DEC-2004;
Bayar HealthCare AG (DE)	
FEATURES	
source	Location/Qualifiers
	1..3407
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
ORIGIN	/db_xref="taxon:9606"
Query Match	100.0%; Score 2184; DB 6; Length 3407;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2184; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 AGTCGCAAACTTACACTCTAACTGATTACTTAAAAATACTTATAGACTGAAGTTATAC 60
Db	190 AGTCGCAAACTTACACTCTAACTGATTACTTAAAAATACTTATAGACTGAAGTTATAC 249
Qy	61 TCCTTAGATCGATTTCAGATCATGATATCTCTACAAACAAGAAAATAATATCTTGTTGA 120
Db	250 TCCTTAGATCGATTTCAGATCATGATATCTCTACAAACAAGAAAATAATATCTTGTTGA 309
Qy	121 TTCAATGCTGAATATGGAAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
Db	310 TTCAATGCTGAATATGGAAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT 369
Qy	181 GGACATCTCATCAATGATTAATTCATATCTCTGATGGCGAGTTTATTTCTCTTAGAATAC 240
Db	370 GGACATCTCATCAATGATTAATTCATATCTCTGATGGCGAGTTTATTTCTCTTAGAATAC 429
Qy	241 AACTACGTGAAGCAATGGAGCATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300
Db	430 AACTACGTGAAGCAATGGAGCATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 489
Qy	301 AAAAGGCGAGCTGATTACAGAGAGAGGAGTTCCAAACAACACACAGTGGGTCCACATGGTCA 360
Db	490 AAAAGGCGAGCTGATTACAGAGAGAGGAGTTCCAAACAACACACAGTGGGTCCACATGGTCA 549
Qy	361 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATTTGAACCA 420
Db	550 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATTTGAACCA 609
Qy	421 AATTATCCAGTTACAGAAATCACATGGACGGGGAAGAAGATATAATATATAATGGAATA 480
Db	610 AATTATCCAGTTACAGAAATCACATGGACGGGGAAGAAGATATAATATATAATGGAATA 669
Qy	481 ACTGACTGGGTTTATGAGAGGAGTCTTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCCA 540
Db	670 ACTGACTGGGTTTATGAGAGGAGTCTTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCCA 729
Qy	541 AACGGGCACTTTTTTAGCATATATGCCCAATTTAAACGACACAGAAATGCCCACTTTATTGAATAC 600
Db	730 AACGGGCACTTTTTTAGCATATATGCCCAATTTAAACGACACAGAAATGCCCACTTTATTGAATAC 789
Qy	601 TCCTTTCTACTCTGATGAGTCACTGCAGTACCCCAAGACTGTACGGGTTCCATATCCAAAG 660
Db	790 TCCTTTCTACTCTGATGAGTCACTGCAGTACCCCAAGACTGTACGGGTTCCATATCCAAAG 849
Qy	661 GCAGGAGCTGTGAATCCAACTGTAAGTCTTTTGTGTGTAATATACAGACTCTCTCAGTCTCA 720
Db	850 GCAGGAGCTGTGAATCCAACTGTAAGTCTTTTGTGTGTAATATACAGACTCTCTCAGTCTCA 909
Qy	721 GTCAACCAATGCAATCTTCATACAAATCACTGCTCTCTGCTCTATGTTGATAGGGGATCAC 780
Db	910 GTCAACCAATGCAATCTTCATACAAATCACTGCTCTCTGCTCTATGTTGATAGGGGATCAC 969
Qy	781 TACTTGTGTGATGTGACATGGCGCAACACAAGAAAAGAAATTTCTTTCCAGTGGCTCAGGAGG 840

Db	1630	TTGATGAAACAAATTTTGGTATCAGATGATCTTGCTCCTCAATTTTGTATAAATCCAG	1680
Qy	1501	AAATATCCTCTACTATTAGATGTGTATGCAGGCCCATGTAGTCAAAAAGCAGACACTGTCT	1560
Db	1690	AAATATCCTCTACTATTAGATGTGTATGCAGGCCCATGTAGTCAAAAAGCAGACACTGTCT	1749
Qy	1561	TTCCAGACTGAACCTGGGCCACTTACCTTTGCAGGACACAGAAAACATTTAGTAGCTAGCTTT	1620
Db	1750	TTCCAGACTGAACCTGGGCCACTTACCTTTGCAGGACACAGAAAACATTTAGTAGCTAGCTTT	1809
Qy	1621	GATGGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCANTGCAATCAACAGAAAGACTG	1680
Db	1810	GATGGCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGCANTGCAATCAACAGAAAGACTG	1869
Qy	1681	GGAACTTTTGAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT	1740
Db	1870	GGAACTTTTGAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT	1929
Qy	1741	GTGCACACAAACGAATTTGCCAATTTGGGGCTGGTTCATATGGAGGTAAGTAACCTCAATG	1800
Db	1930	GTGCACACAAACGAATTTGCCAATTTGGGGCTGGTTCATATGGAGGTAAGTAACCTCAATG	1989
Qy	1801	GTCTCTGGGATCGGGAAGTGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCCTGTATCCCGG	1860
Db	1990	GTCTCTGGGATCGGGAAGTGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCCTGTATCCCGG	2049
Qy	1861	TGGGAGTACTATGACTCAGTGTACACAGAAAGTTTACATGGGTCTCCCAACTCCAGAAAGAC	1920
Db	2050	TGGGAGTACTATGACTCAGTGTACACAGAAAGTTTACATGGGTCTCCCAACTCCAGAAAGAC	2109
Qy	1921	AACCTTGACCAATTACAGAAATTTCAACAGTCAATGAGCAGAGCTGAAAATTTTAAACAAGTT	1980
Db	2110	AACCTTGACCAATTACAGAAATTTCAACAGTCAATGAGCAGAGCTGAAAATTTTAAACAAGTT	2169
Qy	1981	GAGTACCTTCCTTATTCATGGAACAGCAGATGATAACGTTCACTTTCAGCAGTCAGCTCAG	2040
Db	2170	GAGTACCTTCCTTATTCATGGAACAGCAGATGATAACGTTCACTTTCAGCAGTCAGCTCAG	2229
Qy	2041	ATCTCCAAAGCCCTGGTGCAGTGTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA	2100
Db	2230	ATCTCCAAAGCCCTGGTGCAGTGTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA	2289
Qy	2101	GACCATGGATAGCTAGCAGCAGCAGCACAACAATATATATATATATATATATATATATATAT	2160
Db	2290	GACCATGGATAGCTAGCAGCAGCAGCACAACAATATATATATATATATATATATATATATAT	2349
Qy	2161	ATAAAACAATGTTTCTCTTTACCT	2184
Db	2350	ATAAAACAATGTTTCTCTTTACCT	2373
RESULT 4			
LOCUS	AR653630	3407 bp	DNA linear PAT 13-JUN-2005
DEFINITION	Sequence 5 from patent US 6867679.		
ACCESSION	AR653630		
VERSION	AR653630.1	GI:67584135	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
AUTHORS	Brown,N.J.		
TITLE	Methods for identifying contraindications to angiotensin converting enzyme inhibitor and/or vasopetidase inhibitor treatment		
JOURNAL	Patent: US 6867679-A 5 03-MAY-2005;		
FEATURES	vanderbilt University; Nashville, TN		
source	Location/Qualifiers		
	1. 3407		
	/organism="unknown"		
ORIGIN	/mol_type="genomic DNA"		

[illegible]

Db 2290 GACATGGGAATAGCTAGCAGCAGCACACCAACATATATATATACCCACATGACCCACTTC 2344

Qy 2161 ATAAACAATGTTTCTCTTTTACCT 2184
 |||
 2350 ATAAACAATGTTTCTCTTTACCT 2373

Db

RESULT 5
 HSPCHDP7
 LOCUS Human pCHDP7 mRNA for liver dipeptidyl peptidase IV. 3407 bp mRNA linear PRI 16-NOV-1993
 DEFINITION
 ACCESSION X60708 S40353
 VERSION X60708.1 GI:35335
 KEYWORDS dipeptidyl peptidase IV; pCHDP7.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Butheria; Euarchontoglires; Primates; Catarrhini;
 Hominiidae; Homo.

REFERENCE
 1 Misumi, Y., Hayashi, Y., Arakawa, P. and Ikehara, Y.
 AUTHORS Molecular cloning and sequence analysis of human dipeptidyl
 TITLE peptidase IV, a serine proteinase on the cell surface
 JOURNAL Biochim. Biophys. Acta 1131 (3), 333-336 (1992)
 PUBMED 1352704
 REFERENCE 2 (bases 1 to 3407)
 AUTHORS Misumi, Y.
 TITLE Direct Submission
 JOURNAL Submitted (30-JUN-1991) Y. Misumi, Second Dept of Biochemistry,
 School of Medicine, Fukuoka University, 7-45-1 Nanakuma Jonnan-ku,
 Fukuoka 814-01, JAPAN
 COMMENT On Aug 11, 2005 this sequence version replaced gi:251771.
 See Ann. Hum. Genet. 54:191-197(1990) for overlapping sequence.

FEATURES
 source
 1..3407
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="liver"
 /clone_lib="lambda ZAPII"
 1..3407
 /gene="pCHDP7"
 1..3407
 /gene="pCHDP7"
 /evidence=experimental
 76..2376
 /gene="pCHDP7"
 /EC number="3.4.14.5"
 /codon_start=1
 /product="dipeptidyl peptidase iv"
 /protein_id="CAA43118.1"
 /db_xref="GI:35336"
 /db_xref="GON:P27487"
 /db_xref="UniProt/Swiss-Prot:P27487"
 /translation="MKTPTWKILGLGAAALVITITVPVLLNKGTDATADSRKRTYT
 LTDLKNYVRLKLYLSRWTSDEHYLYKQNNILVFNAEGNSVFLNENSTFDFGHSI
 NDYSIPDQGFILLEYVYKQWHSYTSYDIVLKNROLITEERIPNNTQWTFSPV
 GKHLATVANDIYVKLEPNLPSYRIITGKEDIIYNGIDTWYEEVFSAYSALWSP
 NGFLAYQNDITEVPLIYSFYSDESLOTPKTRVPIYPRAGAVNPTVRFVNTUSL
 SVYNTATSIQTAPASMLIGDHYLCDVTWATQERISLQWLRITQNTYSVMDICDYDESS
 GRWNCILVARQHIEMTSTGWGFRPSPPEHFTLDGNSFYKIIISNEEGRHICYPIDKK
 DCTFITKGTWYFIEGIALTSDLYLISNEYKGMPPGRNLVKIOLIDYTKVTCLSCLN
 PERCOYYSVFSEAKYQDLRGSGPGLPLYTLHSSVNDKGLRVLENSALDKMLQNVQ
 MFSKLDLFIILNETKFWYQMLIPPHFDKSKYPLLLDDVYAGPCSQADVTVRLNWTY
 LASTENIIVASFDGRSGSGQDKIMHAINRRLGTPEVDQIEAARQFSKMGFFVNDKRI
 ALMGWGYGVYTSVLMGSGSGYGFVKGIAPVAPSVYTERYMGLPFPEDNLDRI
 YRNSTVSAENFKQVEYLLIHGTADDNVHFQQAISKALVDVGVDVFOAMWYTDDEH
 GIASSTAQHIIYTHMSHFIKQCFSLP"

ORIGIN

Query Match 100.0%; Score 2184; DB 8; Length 3407;
 Best local similarity 100.0%; Pred. No. 0;
 Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

[illegible]

Db 1571 GCTTTGGATAAAATGCTGCAGAAATGTCCAGATGCCCTCCAAAAAATCGGACTTCATTATT 1630
QY 1441 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGCTCCTCATTTTGATATAATCCAAAG 1500
Db 1631 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGCTCCTCATTTTGATATAATCCAAAG 1690
QY 1501 AAATATCCTCTACTATTAGATGTGTATGAGGCGCCATGTAGTCAAAAAGCAGACACTGTC 1560
Db 1691 AAATATCCTCTACTATTAGATGTGTATGAGGCGCCATGTAGTCAAAAAGCAGACACTGTC 1750
QY 1561 TTCAGATGAACTGGGCCACTTACCTTGAAGCAGACAGAAAACATTATAGTAGCTAGCTTT 1620
Db 1751 TTCAGACTGAACTGGGCCACTTACCTTGAAGCAGACAGAAAACATTATAGTAGCTAGCTTT 1810
QY 1621 GATGCGAGGAAGTGGTTACCAAGGAGATAGATCATGTCATCAACAGAGACTG 1680
Db 1811 GATGCGAGGAAGTGGTTACCAAGGAGATAGATCATGTCATCAACAGAGACTG 1870
QY 1681 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAATTTTCAAAAATGGGATTT 1740
Db 1871 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAATTTTCAAAAATGGGATTT 1930
QY 1741 GTGGACAAACAAACGAATTTGCGGCTGGTCAATATGGAGGGTACGTAAACCTCAATG 1800
Db 1931 GTGGACAAACAAACGAATTTGCGGCTGGTCAATATGGAGGGTACGTAAACCTCAATG 1990
QY 1801 GTCTGGGATCGGGAAGTGGCGTTCATGATGTGGATAGCCGTGGCGCTGTATCCCGG 1860
Db 1991 GTCTGGGATCGGGAAGTGGCGTTCATGATGTGGATAGCCGTGGCGCTGTATCCCGG 2050
QY 1861 TGGGAGTACTATGACTCAGGTGTACACAGAACTGTACATGGGTCTCCCAACTCCAGAAGAC 1920
Db 2051 TGGGAGTACTATGACTCAGGTGTACACAGAACTGTACATGGGTCTCCCAACTCCAGAAGAC 2110
QY 1921 AACCTTGACCATTTACAGAAATTTCAACAGTCTATGAGCAGAGCTGAAAATTTTAAACAAGTT 1980
Db 2111 AACCTTGACCATTTACAGAAATTTCAACAGTCTATGAGCAGAGCTGAAAATTTTAAACAAGTT 2170
QY 1981 GAGTACCTCTTATTCATGGAACAGCAGATGATTAAGTTTCACTTTCAGCAGTCAAGTCAG 2040
Db 2171 GAGTACCTCTTATTCATGGAACAGCAGATGATTAAGTTTCACTTTCAGCAGTCAAGTCAG 2230
QY 2041 ATCTCCAAAGCCCTGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGA 2100
Db 2231 ATCTCCAAAGCCCTGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGA 2290
QY 2101 GACCATGGAATAGCTAGCAGCAGCAGCACCAACATATATATACCCACATGAGCCACTTC 2160
Db 2291 GACCATGGAATAGCTAGCAGCAGCAGCACCAACATATATATATACCCACATGAGCCACTTC 2350
QY 2161 ATAAAACAATGTTTCTTTACCT 2184
Db 2351 ATAAAACAATGTTTCTTTACCT 2374

RESULT 9
CQ776356
LOCUS 3445 bp DNA linear PAT 11-MAR-2004
DEFINITION Sequence 42 from Patent EP1394274.
ACCESSION CQ776356
VERSION CQ776356.1 GI:45379746
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuwara,K.
AUTHORS Methods of testing for bronchial asthma or chronic obstructive
TITLE pulmonary disease
Patent: EP 1394274-A 42 03-MAR-2004;
JOURNAL Genox Research, Inc. (JP)

FEATURES
source Location/Qualifiers
1..3445
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.9%; Score 2182.4; DB 6; Length 3445;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGTCGCAAAACCTTACACTCTAACTGATTACTTTAAATAATCTATAGACTGAAGTTATAC 60
Db 225 AGTCGCAAAACCTTACACTCTAACTGATTACTTTAAATAATCTATAGACTGAAGTTATAC 284
QY 61 TCCTTAAAGATGGATTCAGATCATGAATATCTCTACAAACAGAAAATAATCTTGGTA 120
Db 285 TCCTTAAAGATGGATTCAGATCATGAATATCTCTACAAACAGAAAATAATCTTGGTA 344
QY 121 TTCAATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
Db 345 TTCAATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 404
QY 181 GGACATTTCTATCAATGATTATTCAATATCTCTGATGGGAGTTTATTCTCTTAGAATAC 240
Db 405 GGACATTTCTATCAATGATTATTCAATATCTCTGATGGGAGTTTATTCTCTTAGAATAC 464
QY 241 AACTAGGTGAAGCAATGGAGGATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300
Db 465 AACTAGGTGAAGCAATGGAGGATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 524
QY 301 AAAAGGCGAGCTGATTACAGAGAGAGAGGATTCCAAACAAACACACAGTGGGTCACTGGTCA 360
Db 525 AAAAGGCGAGCTGATTACAGAGAGAGAGGATTCCAAACAAACACACAGTGGGTCACTGGTCA 584
QY 361 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATGAAACA 420
Db 595 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATGAAACA 644
QY 421 AATTACCAAGTTACAGAAATCAGATGAGAGGGGAAAGATATATAATATATGGAATA 480
Db 645 AATTACCAAGTTACAGAAATCAGATGAGAGGGGAAAGATATATAATATATGGAATA 704
QY 481 ACTGACTGGGTTTATGAAGAGAGAGTCTTCACTGCTCTCTGCTCTGTGGTGGTCTCCA 540
Db 705 ACTGACTGGGTTTATGAAGAGAGAGTCTTCACTGCTCTCTGCTCTGTGGTGGTCTCCA 764
QY 541 AACGGCAGCTTTTATGACATATGCCCAATTTAAACGACACAGAAAGTCCCAGTTTATGAAATAC 600
Db 765 AACGGCAGCTTTTATGACATATGCCCAATTTAAACGACACAGAAAGTCCCAGTTTATGAAATAC 824
QY 601 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTCCCATATCCAAAG 660
Db 825 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTCCCATATCCAAAG 884
QY 661 GCAGGAGCTGGAATCCCAACTGTAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA 720
Db 885 GCAGGAGCTGGAATCCCAACTGTAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA 944
QY 721 GTCCACCAATGCAACTTCCATACAAATCACTGCTCTGCTCTTATGTTGATAGGGATCAC 780
Db 945 GTCCACCAATGCAACTTCCATACAAATCACTGCTCTGCTCTTATGTTGATAGGGATCAC 1004
QY 781 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCGAGTGGCTCAGGAGG 840
Db 1005 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCGAGTGGCTCAGGAGG 1064
QY 841 ATTTCAGAACTATTTCGGTTCATGATATTTGTGACTATGATGAATCCAGTGGAGATGGAAC 900
Db 1065 ATTTCAGAACTATTTCGGTTCATGATATTTGTGACTATGATGAATCCAGTGGAGATGGAAC 1124
QY 901 TGCTTAGTGGCACGGCAACACATTTGAAATGAGTACTACTGCTGGGTGGGAAGATTTAGG 960

[illegible]

Db	1725	AAATATCCTCTACTATTAGATGTGTAATGCAGGCCCAATGTAGTCAAAGACGACACTGTC	1780
Qy	1561	TTCCAGACTGAACCTGGGCCCACTTACCCTTTGCAAGCACAGAAAAACAATTATFAGTAGCTAGCTTT	1620
Db	1785	TTCCAGACTGNACTGGGCCCACTTACCCTTGCAAGCACAGAAAAACAATTATFAGTAGCTAGCTTT	1844
Qy	1621	GATGGCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGCAATGCAATCAAAGAAGACTG	1680
Db	1845	GATGGCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGCAATGCAATCAAAGAAGACTG	1904
Qy	1681	GGAACATTTGAAAGTTGAAGTCAAAATGAAGCAGCGACAGACAATTTTCAAAAATGGGATTT	1740
Db	1905	GGAACATTTGAAAGTTGAAGTCAAAATGAAGCAGCGACAGACAATTTTCAAAAATGGGATTT	1964
Qy	1741	GTGGACAACAAACGAATTGGCAATTTGGGGCTGGTCAATATGGAGGGTACGTAACCTCAATG	1800
Db	1965	GTGGACAACAAACGAATTTGGCAATTTGGGGCTGGTCAATATGGAGGGTACGTAACCTCAATG	2024
Qy	1801	GTCCCTGGGATCGGGAAATGGCGTTTCAAGTGTGSAATAGCCGTGGCGCTGTATCCCAG	1860
Db	2025	GTCCCTGGGATCGGGAAATGGCGTTTCAAGTGTGSAATAGCCGTGGCGCTGTATCCCAG	2084
Qy	1861	TGGGAGTACTATGACTCAGTGTACACAGAACGTTTACATGGGTCTCCCNACTCCAGAGAC	1920
Db	2085	TGGGAGTACTATGACTCAGTGTACACAGAACGTTTACATGGGTCTCCCNACTCCAGAGAC	2144
Qy	1921	AACCTTGACCATTACAGAAATTCACAGTCATGAGCAGAGCTCAAAAATTTTAAACAAGTT	1980
Db	2145	AACCTTGACCATTACAGAAATTCACAGTCATGAGCAGAGCTCAAAAATTTTAAACAAGTT	2204
Qy	1981	GAGTACCTCCTTATTTCATGGAACAGCAGATGATAACGTTTCACTTCAGCAGTCAGCTCAG	2040
Db	2205	GAGTACCTCCTTATTTCATGGAACAGCAGATGATAACGTTTCACTTCAGCAGTCAGCTCAG	2264
Qy	2041	ATCTCCAAGCCCTGTGCATGTTGGAGTGGATTTCCAGGCATATGGTATACCTGATGAA	2100
Db	2265	ATCTCCAAGCCCTGTGCATGTTGGAGTGGATTTCCAGGCATATGGTATACCTGATGAA	2324
Qy	2101	GACCATGGAATAGCTAGCAGCAGACAGACCAACATATATATACCCACATGAGCCACTTC	2160
Db	2325	GACCATGGAATAGCTAGCAGCAGACAGACCAACATATATATACCCACATGAGCCACTTC	2384
Qy	2161	ATAAACAATGTTTCTCTTTACCT	2184
Db	2385	ATAAACAATGTTTCTCTTTACCT	2408
RESULT 11			
CS031460		3445 bp	DNA linear PAT 10-MAR-2005
LOCUS			
DEFINITION		Sequence 966 from Patent WO2005016962.	
ACCESSION		CS031460	
VERSION		CS031460.1	GI:60731491
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
		Homidae; Homo.	
REFERENCE		1	
AUTHORS		Abbas,A., Clark,H., Ouyang,W., Williams,M.P., Wood,W.I. and Wu,T.D.	
TITLE		Compositions and methods for the treatment of immune related	
JOURNAL		Patent: WO 2005016962-A 966 24-FEB-2005;	
FEATURES		Genentech, Inc. (US)	
source		Location/Qualifiers	
		1..3445	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Query Match		99.9%; Score 2182.4; DB 6; Length 3445;	

Best Local Similarity 99.9%; Pred. No. 0; Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	AGTCGCAAAACTTACACTCTAACTGATTA	CTTAAAAAATACTTTATAGACTGAAAGTTTATAC 60
Db	225	AGTCGCAAAACTTACACTCTAACTGATTA	CTTAAAAAATACTTTATAGACTGAAAGTTTATAC 284
Qy	61	TCCTTAAGATGGATTCAGATCATGAATAT	CTCTACAAACAAGAAATAATATCTTTGGTA 120
Db	285	TCCTTAAGATGGATTCAGATCATGAATAT	CTCTACAAACAAGAAATAATATCTTTGGTA 344
Qy	121	TTCAATGCTGAATATGAAACAGCTCAGTTT	CTTTGGAGAACAGTACATTTTGATGAGTTT 180
Db	345	TTCAATGCTGAATATGAAACAGCTCAGTTT	CTTTGGAGAACAGTACATTTTGATGAGTTT 404
Qy	181	GGACATTTCTATCAATGATTAATCAATAT	CTCTGATGGGCAGTTTATTTCTTTAGAATAC 240
Db	405	GGACATTTCTATCAATGATTAATCAATAT	CTCTGATGGGCAGTTTATTTCTTTAGAATAC 464
Qy	241	AATTAACGTAAGCAATGGAGGCAATTCCTA	CAACAGCTTCAATGACATTTTAAAT 300
Db	465	AATTAACGTAAGCAATGGAGGCAATTCCTA	CAACAGCTTCAATGACATTTTAAAT 524
Qy	301	AAAAGGCAGCTGATTACAGAAAGAGAGNT	TCCAAAACAACACAGCTGGGTCAATGGTCA 360
Db	525	AAAAGGCAGCTGATTACAGAAAGAGAGNT	TCCAAAACAACACAGCTGGGTCAATGGTCA 584
Qy	361	CCAGTGGGTCAATAATGGCATATGTTTGGAA	CAATGACATTTATGTTAAATTTGAACCA 420
Db	585	CCAGTGGGTCAATAATGGCATATGTTTGGAA	CAATGACATTTATGTTAAATTTGAACCA 644
Qy	421	AATTTACCAAGTTTACAGAAATCAATGGA	ATCCAGCGGGAAAGAGATATATATATGAATA 480
Db	645	AATTTACCAAGTTTACAGAAATCAATGGA	ATCCAGCGGGAAAGAGATATATATATGAATA 704
Qy	481	ACTGACTGGGTTTATGAAGGGAAGCTTTC	CAGTGCCTACTCTGCTCTGTGGTCTCCA 540
Db	705	ACTGACTGGGTTTATGAAGGGAAGCTTTC	CAGTGCCTACTCTGCTCTGTGGTCTCCA 764
Qy	541	AACGGCACTTTTTTTAGCATATGCCAATTT	AAACGACACAGAACTCCCACTTATTTGAATAC 600
Db	765	AACGGCACTTTTTTTAGCATATGCCAATTT	AAACGACACAGAACTCCCACTTATTTGAATAC 824
Qy	601	TCCTTCTACTCTGATGAGTCACTGCAGTAC	CCCAAAGACTGTACGGGTCCATATCCAAAG 660
Db	825	TCCTTCTACTCTGATGAGTCACTGCAGTAC	CCCAAAGACTGTACGGGTCCATATCCAAAG 884
Qy	661	GCAGGAGCTGTGAATCCAACTGTAAAGTT	CTTTTGTGTAATAACAGACTCTCTCAGCTCA 720
Db	885	GCAGGAGCTGTGAATCCAACTGTAAAGTT	CTTTTGTGTAATAACAGACTCTCTCAGCTCA 944
Qy	721	GTCAACCAATGCAACTTCCATACAAATCA	CTGCTCCTCTCTATGTTGATGGGATCAC 780
Db	945	GTCAACCAATGCAACTTCCATACAAATCA	CTGCTCCTCTCTATGTTGATGGGATCAC 1004
Qy	781	TACTTGTGTGATGTGACATGGGCACACAA	GAAAGAAATTTCTTTGCAAGTGGCTCAGGAGG 840
Db	1005	TACTTGTGTGATGTGACATGGGCACACAA	GAAAGAAATTTCTTTGCAAGTGGCTCAGGAGG 1064
Qy	841	ATTTCAGAACTATTTCGGTCTATGGATAT	TTTGTGACTATGATGAATCCAGTGGAAAGATGGAAC 900
Db	1065	ATTTCAGAACTATTTCGGTCTATGGATAT	TTTGTGACTATGATGAATCCAGTGGAAAGATGGAAC 1124
Qy	901	TGCTTAGTGGCAGGCGACACATTTGAATG	AGTACTGCTGGCTGGGTGGAGATTTAGG 960
Db	1125	TGCTTAGTGGCAGGCGACACATTTGAATG	AGTACTGCTGGCTGGGTGGAGATTTAGG 1184
Qy	961	CTTTTCAGAACTCAATTTTACCTTTGATGG	TAAATAGCTTCTTACAAGATCATCAGCAATGAA 1020
Db	1185	CTTTTCAGAACTCAATTTTACCTTTGATGG	TAAATAGCTTCTTACAAGATCATCAGCAATGAA 1244
Qy	1021	GAAGGTTACAGACACATTTTGCTATTTT	CCAAAATAGATAAAAAAGACTGCACATTTTATTA 1080

Qy	1691	GGAACATTTGAAAGTTGAAGATCAAAATTGAAGCAGCCAGACAATTTTCAAATAATGGGATTT	1741
Db	1905	GGAAATTTGAAAGTTGAAGATCAAAATTGAAGCAGCCAGACAATTTTCAAATAATGGGATTT	1964
Qy	1741	GTGACAAACAAACGAAATTGCAAATTTGGGGCTGGTCATATGAGAGGTACGTAACCTCAATG	1800
Db	1965	GTGACAAACAAACGAAATTGCAAATTTGGGGCTGGTCATATGAGAGGTACGTAACCTCAATG	2024
Qy	1801	GTCCTGGGATCGGAAAGTGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCCTGTATCCCGG	1860
Db	2025	GTCCTGGGATCGGAAAGTGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCCTGTATCCCGG	2084
Qy	1861	TGGAGTACTATGACTCAGTGTACACAGAAGCTTACATGGGTCTCCCACTCCAGAAGAC	1920
Db	2085	TGGAGTACTATGACTCAGTGTACACAGAAGCTTACATGGGTCTCCCACTCCAGAAGAC	2144
Qy	1921	AACCTTGACCATTACAGAAATTCACAGTCTANTAGCAGAGCTGAAAAATTTTTAAACAAGTT	1980
Db	2145	AACCTTGACCATTACAGAAATTCACAGTCTANTAGCAGAGCTGAAAAATTTTTAAACAAGTT	2204
Qy	1981	GAGTACCTCTCTATTTCATGGAACAGCAGATGATAACGTTCACTTTACGAGTCAGCTCAG	2040
Db	2205	GAGTACCTCTCTATTTCATGGAACAGCAGATGATAACGTTCACTTTACGAGTCAGCTCAG	2264
Qy	2041	ATCTCCAAAAGCCCTGGTGCATGTTGGAGTGTGATTTCCAGGCAATGTCGTATACGTATGA	2100
Db	2265	ATCTCCAAAAGCCCTGGTGCATGTTGGAGTGTGATTTCCAGGCAATGTCGTATACGTATGA	2324
Qy	2101	GACCATGGATAGCTACGCAGCACACGCCACATATATATACCCACATGAGCCACTTC	2160
Db	2325	GACCATGGATAGCTACGCAGCACACGCCACATATATATACCCACATGAGCCACTTC	2384
Qy	2161	ATAAAACAAATGTTTCTCTTTACCT	2184
Db	2385	ATAAACAATGTTTCTCTTTACCT	2408
RESULT 13			
CS040412			
LOCUS	CS040412	3445 bp	DNA linear PAT 22-MAR-2005
DEFINITION	Sequence 966 from Patent WO2005019258.		
ACCESSION	CS040412		
VERSION	CS040412.1	GI:61848176	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Abbas, A., Clark, H., Ouyang, W., Williams, P.M., Wood, W.I. and Wu, T.D.		
TITLE	Compositions and methods for the treatment of immune related diseases		
JOURNAL	Patent: WO 2005019258-A 966 03-MAR-2005;		
	Genentech, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..3445		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	99.9%	Score 2182.4;	DB 6; Length 3445;
Best Local Similarity	99.9%	Pred. No. 0;	
Matches 2183;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	AGTCGCAAACTTACACTCTAACCTGATTACTTAAAAAATACTTATAGACTGAAGTTATAC	60
Db	225	AGTCGCAAACTTACACTCTAACCTGATTACTTAAAAAATACTTATAGACTGAAGTTATAC	284
Qy	61	TCCTTAGATGGATTTCAGATCATGAATATCTCTTACAAAACAAGAAAATAATATCTTGGTA	120

QY	1201	TATACAAAGTGCATGCTCAGTTGTGAGCTGAATCCGGAAGGTGTCAGTACTATTCT	1260
Db	1425	TATACAAAGTGCATGCTCAGTTGTGAGCTGAATCCGGAAGGTGTCAGTACTATTCT	1484
QY	1261	GTGTCAVTCAGTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCGTCTGCC	1320
Db	1485	GTGTCAVTCAGTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCGTCTGCC	1544
QY	1321	CTCTATACCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGCAATTC	1380
Db	1545	CTCTATACCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGCAATTC	1604
QY	1381	GCTTTGGATAAATCTGTCAGATGTCAGATGCCCTCCAAAACACTGGACTTCATTATT	1440
Db	1605	GCTTTGGATAAATCTGTCAGATGTCAGATGCCCTCCAAAACACTGGACTTCATTATT	1664
QY	1441	TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCATTTTGTATAAATCCAAG	1500
Db	1665	TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCATTTTGTATAAATCCAAG	1724
QY	1501	AAATATCTCTATCTATTAGATGTGTATGAGGCCCATGTAGTCAAAAAGCAGACATGTC	1560
Db	1725	AAATATCTCTATCTATTAGATGTGTATGAGGCCCATGTAGTCAAAAAGCAGACATGTC	1784
QY	1561	TTCAGACTGAATGGGCCACTTTACCTTGCAAGCAGACAGAAAACATTATAGTAGCTT	1620
Db	1785	TTCAGACTGAATGGGCCACTTTACCTTGCAAGCAGACAGAAAACATTATAGTAGCTT	1844
QY	1621	GATGCGAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGACTG	1680
Db	1845	GATGCGAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGACTG	1904
QY	1681	GGAACTTTGAAGTTGAAGATCAAAATGGAAGCAGCAGACAATTTTCAAAAATGGGATTT	1740
Db	1905	GGAACTTTGAAGTTGAAGATCAAAATGGAAGCAGCAGACAATTTTCAAAAATGGGATTT	1964
QY	1741	GTGGACACAAACGAATTTGCGGCTGGTTCATATGGAGGGTACGTAACTCAATG	1800
Db	1965	GTGGACACAAACGAATTTGCGGCTGGTTCATATGGAGGGTACGTAACTCAATG	2024
QY	1801	GTCTGGGATCGGGAAGTGGCGTGTTCAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	1860
Db	2025	GTCTGGGATCGGGAAGTGGCGTGTTCAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	2084
QY	1861	TGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAGAAGAC	1920
Db	2085	TGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAGAAGAC	2144
QY	1921	AACTTTGACCATTTACAGAAATTTCAACAGTCATGCGAGAGCTGAAATTTTAAACAGTT	1980
Db	2145	AACTTTGACCATTTACAGAAATTTCAACAGTCATGCGAGAGCTGAAATTTTAAACAGTT	2204
QY	1981	GAGTACCTCTTATTATGGAACAGCAGATGATTAAGTTTCACTTTTCAGCAGTCAGCTCAG	2040
Db	2205	GAGTACCTCTTATTATGGAACAGCAGATGATTAAGTTTCACTTTTCAGCAGTCAGCTCAG	2264
QY	2041	ATCTCCAAAGCCCTGGTGCATGTTGGAGTGAATTTCCAGGCAATGTGGTATATGATGAA	2100
Db	2265	ATCTCCAAAGCCCTGGTGCATGTTGGAGTGAATTTCCAGGCAATGTGGTATATGATGAA	2324
QY	2101	GACCATGGATAGCTAGCAGCAGCAGCACCAACATATATATATATATATATATATATAT	2160
Db	2325	GACCATGGATAGCTAGCAGCAGCAGCACCAACATATATATATATATATATATATATAT	2384
QY	2161	ATAAAAAATGTTTCTCTTACCT 2184	
Db	2385	ATAAAAAATGTTTCTCTTACCT 2408	

RESULT 14
CS041831
LOCUS

CS041831 3445 bp DNA linear PAT 22-MAR-2005

Sequence 2385 from Patent WO2005019258.	
CS041831	
CS041831.1	GI:61849188
Homo sapiens (human)	
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
1	
Abbas, A., Clark, H., Ouyang, W., Williams, P.M., Wood, W.I. and Wu, T.D.	
Compositions and methods for the treatment of immune related diseases	
Patent: WO 2005019258-A 2385 03-MAR-2005;	
Genentech, Inc. (US)	
Location/Qualifiers	
1..3445	
/organism="Homo sapiens"	
/mol_type="unassigned DNA"	
/db_xref="taxon:9606"	
Query Match	99.9%; Score 2182.4; DB 6; Length 3445;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 2183; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	1 AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATCTTATAGACTGAAGTTATAC 60
Db	225 AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATCTTATAGACTGAAGTTATAC 284
QY	61 TCCTTAAGATGGATTTTCAAGATCATGAATATCTCTACAAAACAGAAAAATATATCTTTGGTA 120
Db	285 TCCTTAAGATGGATTTTCAAGATCATGAATATCTCTACAAAACAGAAAAATATATCTTTGGTA 344
QY	121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
Db	345 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 404
QY	181 GGACATTTCTATCAATGATTATTCATATCTCTGATGGCGAGTTTATCTCTTAGAATAC 240
Db	405 GGACATTTCTATCAATGATTATTCATATCTCTGATGGCGAGTTTATCTCTTAGAATAC 464
QY	241 AACTACGTGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300
Db	465 AACTACGTGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 524
QY	301 AAAAGGCACTGATTTACAGAGAGAGGATTCACAAACACACAGTGGGTCACTGGTCA 360
Db	525 AAAAGGCACTGATTTACAGAGAGAGGATTTCCAAACACACACAGTGGGTCACTGGTCA 584
QY	361 CCAGTGGGTCAATAATTTGGCATATGTTTGGACATGATTTATGTTAAAAATTCAGACCA 420
Db	585 CCAGTGGGTCAATAATTTGGCATATGTTTGGACATGATTTATGTTAAAAATTCAGACCA 644
QY	421 AATTTCACCAAGTTACAGAAATCACATGGACGGGAAAGAGATATATATATATGGAATA 480
Db	645 AATTTCACCAAGTTACAGAAATCACATGGACGGGAAAGAGATATATATATATGGAATA 704
QY	481 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA 540
Db	705 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA 764
QY	541 AACGGCACTTTTATGAGCATATGCCCAATTTTAAACGACACAGAGTCCCACCTTATGGAATAC 600
Db	765 AACGGCACTTTTATGAGCATATGCCCAATTTTAAACGACACAGAGTCCCACCTTATGGAATAC 824
QY	601 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCATATCCAAAG 660
Db	825 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCATATCCAAAG 884
QY	661 GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTTGTGTAAATATACAGACTCTCTCAGCTCA 720
Db	885 GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTTGTGTAAATATACAGACTCTCTCAGCTCA 944

Search completed: February 17, 2006, 07:42:14
Job time : 11148 secs

THIS PAGE BLANK (user)